9-2009

Investigation of Chemotaxis Genes and Their Functions in Geobacter Species

Hoa T. Tran
University of Massachusetts Amherst, hoatranleeds@yahoo.com

INVESTIGATION OF CHEMOTAXIS GENES AND THEIR FUNCTIONS IN GEOBACTER SPECIES

A Dissertation Presented

by

HOA T. TRAN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2009

Department of Chemistry

INVESTIGATION OF CHEMOTAXIS GENES AND THEIR FUNCTIONS IN GEOBACTER SPECIES

A Dissertation Presented

by

HOA T. TRAN

Approved as to style and content by:

_____
Robert M. Weis, Chair

_____
Derek R. Lovley, Co-chair

_____
Michael J. Knapp, Member

_____
Michael J. Maroney, Member

_____
Bret Jackson, Department Head
Department of Chemistry

DEDICATION

To my loving husband, my family, and my brother Ngoc Tran

ACKNOWLEDGMENTS

The success of this research is due to contributions from many people and I would like to express my gratitude to them. I would like to thank my advisors Dr. Robert M. Weis and Dr. Derek R. Lovley for great support, guidance, and patience, and for giving me freedom to develop and conduct my research. I would like to extend my thanks to Dr. Michael J. Knapp and Dr. Michael J. Maroney for their time serving as committee members during my doctoral studies.

I would like to thanks all my current and former labmates in the Lovley and Weis labs for keeping me in good spirits while I am in the lab, and for their willingness to offer help when I needed it. Special thanks to Frances M. Antommattei and Richard Glaven for their patient mentorship and helpful advice, to Zara Summers for help with microarray analysis, and to Ned Young for help with phylogenetic trees. I would like to express my thanks to Tünde Mester and Xinlei Qian for their advice on protein work. Special gratitude to Lubna Al-Challah, Muktak Aklujkar, and Tatiana Y. Besschetnova for their endless support and great friendship.

I would like to thank my collaborator Julia Krushkal for her help in analyzing the regulation of chemotaxis genes and for making my manuscript readable. I am grateful to John Leszyk and Mary Lipton for performing mass spectrometry analysis. My special thanks to Dale Callaham for helps and advises on microscopic works.

My special thanks to Betsy Blunt and Joy Ward for keeping the lab and my work running smoothly. I also would like to thank the personnel and faculty of the Department of Chemistry and Department of Microbiology of the University of Massachusetts Amherst.

ABSTRACT

INVESTIGATION OF CHEMOTAXIS GENES AND THEIR FUNCTIONS IN GEOBACTER SPECIES

SEPTEMBER 2009

HOA T. TRAN, B.Sc., HANOI UNIVERSITY OF EDUCATION

M.Ed., VICTORIA UNIVERSITY

M. Sc., UNIVERSITY OF LEEDS

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Robert M. Weis and Derek R. Lovley

Geobacter species are δ-*Proteobacteria* and are often predominant in the Fe(III) reduction zone of sedimentary environments. Their abilities to remediate contaminated environments and to produce electricity have inspired extensive studies. Cell motility, biofilm formation, and type IV pili, which have been shown to be regulated by chemotaxis genes in other bacteria, all appear important for the growth of *Geobacter* species in changing environments and for electricity production. The genomes of *Geobacter* species show the presence of a significant number of chemotaxis gene homologs, suggesting important roles for them in the physiology of *Geobacter* species, although gene functions are not yet identified. In this study, we focus on identifying chemotaxis components and studying their functions in *Geobacter* species.

We identified a large number of homologs of chemotaxis genes, which are arranged in six or more major clusters in the genomes of *Geobacter sulfurreducens, Geobacter metallireducens*, and *Geobacter uraniireducens*. Based on homology to known pathways, functions of some chemotaxis clusters were assigned; others appear to

be unique to *Geobacter* species. We discuss the diversity of chemoreceptors and other signaling proteins as well the regulation of chemotaxis genes in *Geobacter* species.

The functions of chemotaxis genes were studied in *G. sulfurreducens*, whose genome contains ~ 70 chemotaxis gene homologs, arranged in 6 major clusters. These chemotaxis clusters are also found in other *Geobacter* species with similar gene order and high level of gene identity, suggesting that our study in *G. sulfurreducens* could be extrapolated to other *Geobacter* species. We identified the function of the *che5* cluster of *G. sulfureducens* as regulation of the biosynthesis of extracellular materials. We showed that *G. sulfurreducens* KN400 is chemotactic, and that this behavior is flagellum-dependent. Our preliminary data indicated that *G. sulfurreducens* may use the *che1* cluster, which is found exclusively in *Geobacteraceae*, to regulate chemotaxis.

Our studies demonstrated important roles of chemotaxis genes in *Geobacter* physiology and their presence in large numbers could be one of the reasons why *Geobacter* species outcompete other species in bioremediation sites. Further studies are warranted for better understanding of the mechanisms of Che-like pathways and their potential use in optimization of conditions for applications of *Geobacter* species in bioremediation and electricity generation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The chemotaxis system in *Escherichia coli*

Escherichia coli cells swim using thin helical filaments called flagella, rotation of which is driven at the base by a motor embedded in the cell envelope, using energy in the form of the transmembrane proton gradient. The motor can switch between two states: clockwise (CW) and counter clockwise (CCW) rotation. In the CCW rotation state, the filaments of a cell unite as a bundle that propels the cell, and the cell has a smooth trajectory called a run. Upon disruption of the bundle as a result of one or more motors reversing to the CW direction, the uncoordinated rotation of filaments leads to rapid somersaulting of the cell, called a tumble (Blair, 1995). In suspension, unstimulated cells execute a three-dimensional random walk: using two modes of swimming, they run in a straight line for about a second, tumble for a fraction of a second, briefly stop, and then randomly swim in an arbitrarily determined new direction (Berg, 2003, Blair, 1995).

Upon encountering a gradient of chemical attractant or repellent in the environment, motile cells use the so-called chemotaxis to control their swimming behavior, swimming towards higher concentrations of attractants and lower concentrations of repellents. The mechanism of chemotaxis has been studied extensively in *E. coli* genetically and biochemically, providing details of the pathway (Bren & Eisenbach, 2000), and has been used as a paradigm for chemotaxis studies in other bacterial species. There are five membrane-bound chemotaxis receptors, also called methyl-accepting chemotaxis proteins (MCPs), involved in the chemotactic signaling pathway of *E. coli*, and six cytoplasmic proteins: an autophosphorylating histidine kinase

(CheA); a linker protein (CheW); a response regulator (CheY), its phosphatase (CheZ), and a methyltransferase (CheR) and methylesterase (CheB) involved in the reversible methylation/demethylation of chemoreceptors: The detailed mechanism is shown in **Figure 1**, and the mechanism is explained as follows.

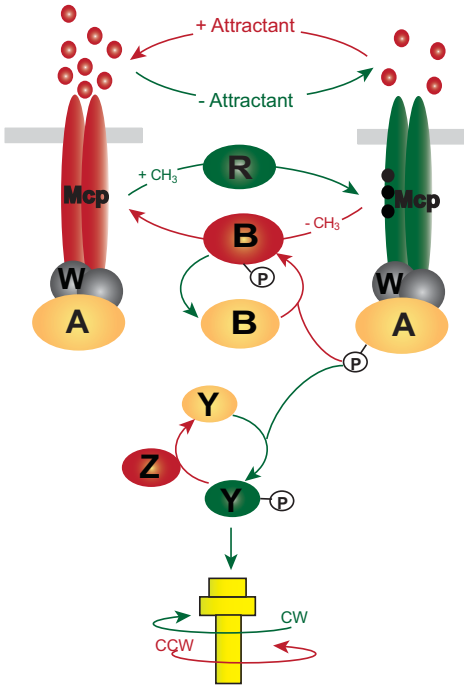


Figure 1. Chemotaxis pathway of *E. coli* (adapted from Parkinson et al. 2005) with explanation in the text

MCPs bind to CheA via CheW to form a receptor complex. Upon binding of an attractant to the periplasmic domain of an MCP, CheA activity is inhibited, and the MCP is methylated on some specific glutamate residues by CheR. When the concentration of attractant decreases (or that of repellent increases), MCPs activate CheA autophosphorylation. Phosphorylated CheA (CheA~P) then transfers its phosphoryl group to CheY that is docked with the receptor complex. Phosphorylated CheY (CheY~P) has lower affinity for the receptor complex than CheY, and thus leaves the complex, diffusing to the motor, where it binds to the FliM protein, switching the motor

to the CW rotation state, leading to a change in cellular motion. The cell changes from a random walking pattern when unstimulated to a tumbling-biased pattern when the concentration of attractant decreases. This signal is terminated quickly (in a fraction of a second), due to the level of CheY~P dropping because of its self-dephosphorylation activity and accelerated dephosphorylation by a phosphatase, CheZ.

In parallel with transfer of a phosphoryl group to CheY, CheA~P also transfers a phosphoryl group to CheB that is docked with the receptor complex, activating this enzyme. CheB~P removes methyl groups from the MCP, which have been previously added by CheR upon binding of the attractant. This reversible methylation/demethylation process brings about the resetting of the chemoreceptor, the so-called adaptation step. Although CheB~P can dephosphorylate itself like CheY~P, CheZ can not hydrolyze CheB~P, and because the status of the chemoreceptor also depends on the activity of CheR, the process of resetting is rather slow. Two processes contribute to dephosphorylation of CheA~P: the fast signaling step that leads to the change in cell behavior (within a fraction of second) and the slow adaptation process (over 3-4 seconds). During this slow adaptation time, the cell has traveled some distance, enabling it to experience changes in attractant or repellent concentration spatially, based on which it makes the decision for its next move: either to continue a run due to higher attractant concentration, or to tumble due to either a lower attractant or higher repellent concentration.

The chemotaxis system is significantly sensitive over a large concentration range of stimuli. It enables cells to detect ~0.1% changes in attractant concentration (Segall *et*

al., 1986, Sourjik & Berg, 2004), and to respond to attractant gradients ranging over 5 orders of magnitude (Adler, 1975).

Chemotaxis-like sytems in other species

Analyses of bacterial genome sequences show that homologs of *E. coli* chemotaxis genes are widespread (Antommattei & Weis, 2006, Galperin, 2005). From these surveys, it is apparent that the MCP and *che* genes in *E. coli* are relatively few in number, which may plausibly reflect modest requirements for sensory transduction in the environment that *E. coli* inhabits. By comparison, the chemotaxis-like systems in other bacteria are greater in number and diversity (Szurmant & Ordal, 2004, Wadhams & Armitage, 2004). Copies of the 'core' genes (*cheAWY*) are clustered in multiple distinct locations and additional genes are present (*cheC, cheD, cheV* and *cheX*) that generate greater mechanistic diversity (Szurmant & Ordal, 2004). For example, Armitage and colleagues have shown that two chemotaxis clusters in the genome of *Rhodobacter sphaeroides* play a role in chemotaxis (Martin *et al.*, 2001), an observation that plausibly reflects the greater need for different signaling pathways in complex environments. Pertinent to the analysis that we present below is the fact that *Geobacter* species also occupy complex ecological niches in sedimentary environments. The published genome of *Geobacter sulfurreducens* has 34 MCP genes and six major *che* gene clusters (Methe *et al.*, 2003); these pathways are likely to play an important role in environmental adaptation.

Biochemical, genetic and physiological investigations of chemotaxis-like signaling pathways in bacteria other than *E. coli* have led to the realization that some of these pathways carry out functions distinct from the well-established role in regulating flagellar motor rotation. These functions include regulation of type IV pilus-dependent

motility, expression of the motility apparatus (both flagella and type IV pili), biofilm formation, and regulation of developmental genes. As examples, *Pseudomonas aeruginosa*, *Rhodospirillum centenum*, *Myxoccocus xanthus*, and *Synechocystis* species all have multiple chemotaxis-like operons that have provided new insight into their diverse functions. *P. aeruginosa* has four major *che* clusters; two are involved in chemotaxis with different suggested roles, a third regulates type IV pilus biogenesis and motility, and the fourth is involved in biofilm formation (Darzins, 1994, Ferrandez *et al.*, 2002, Hickman *et al.*, 2005, Kato *et al.*, 1999, Masduki *et al.*, 1995, Whitchurch *et al.*, 2004). *R. centenum* has three *che* clusters; one mediates chemotaxis, a second regulates cyst development, and a third regulates flagellum biogenesis (Berleman & Bauer, 2005b, Berleman & Bauer, 2005a, Berleman *et al.*, 2004). *M. xanthus* has eight *che* clusters; the functions for four clusters have been identified to date (Zusman *et al.*, 2007). Each cluster regulates a different function, including cell motility, biogenesis of the motility apparatus, and regulation of developmental genes (Blackhart & Zusman, 1985, Bonner *et al.*, 2005, Kirby & Zusman, 2003, Vlamakis *et al.*, 2004, Yang *et al.*, 2000). The functions for two of the three clusters found in the genome of *Synechocystis* PCC6803 have been identified: one regulates type IV pilus-dependent motility, the other pilus biogenesis (Bhaya *et al.*, 2001). Recent studies from Alexandre's group showed the roles of a chemotaxis-like pathway in modulating cell motility, cell-cell-aggregation, and exopolysaccharide production associated with flocculation, as well as cell length of *Azospirillum brasilense* (Bible *et al.*, 2008). As a final example, only one of the three *che* clusters in the *Vibrio cholerae* genome regulates chemotaxis. Mutations in the two

remaining clusters do not affect chemotaxis; their functions are yet to be identified (Butler & Camilli, 2005).

Physiological characteristics and potential applications of Geobacter species

Geobacter species are Gram-negative δ-*Proteobacteria* and are found predominantly in the Fe(III) reduction zone of sedimentary environments. The first *Geobacter* species, *Geobacter metallireducens*, was isolated from freshwater sediments in the Potomac River, Maryland, just downstream from Washington D.C., more than two decades ago (Lovley & Phillips, 1988b). It has a unique metabolism – to extract energy for growth, it can anaerobically oxidize organic compounds completely (to H_2O and CO_2) together with reducing metal compounds. Since then, more *Geobacter* species have been isolated from different types of sediments and characterized, such as *G. sulfurreducens* from surface sediments of a hydrocarbon-contaminated ditch in Norman, Oklahoma (Caccavo *et al.*, 1994b), and *Geobacter uraniireducens* from subsurface sediments undergoing uranium bioremediation (Shelobolina *et al.*, 2008). *Geobacter* species can utilize a variety of organic compounds as sole sources of electron donors coupled with iron compounds as electron acceptors to derive energy. They are able to use a wide range of 'toxic' aromatic compounds as sole electron donors, such as benzene, toluene, phenol and *p*-cresol, and completely oxidize them to CO_2 (Lovley *et al.*, 1993, Lovley *et al.*, 1989), suggesting their potential for bioremediation of organic compounds including aromatic compounds. *Geobacter* species are able to convert some chlorinated compounds to less harmful compounds, suggesting their potential in remediation of chlorinated contaminants in subsurface environments (Strycharz *et al.*, 2008).

Solid waste containing uranium from nuclear power plants, uranium mining, and hospital waste has been a problem of concern. The majority of uranium is in form of U(VI), which is highly soluble and therefore mobile to threaten groundwater. The removal of groundwater contamination is extremely difficult because of technical difficulties and high cost (National Research Council, 1999). Laboratory studies indicated that *Geobacter* species not only could reduce metal compounds such as Fe(III), Mn(IV) and V(V), they could also reduce U(VI) while oxidizing organic compounds. They convert soluble U(VI) to a less soluble form, U(IV), which precipitates as UO_2, suggesting the potential of *Geobacter* species to remediate uranium-contaminated environments. Field tests have been carried out since 2003 at the Old Rifle site in Colorado, a uranium-contaminated aquifer site specified by U.S. Department of Energy, by the injection of acetate (Anderson *et al.*, 2003). Consistent with laboratory data, *Geobacter* species have been found to play a major role in bioremediation of uranium contaminated sites: the concentration of soluble uranium decreased, accompanied by the enrichment of *Geobacter* species in the microbial population (Anderson et al., 2003). Further research has been carried out to improve the remediation process, for which a very promising method comes from the study by Gregory and colleagues, showing that *Geobacter* species can use a graphite electrode as the sole electron donor to reduce uranium (VI) to uranium (IV). In contrast with the traditional approach in which uranium (IV) remains in the environment and poses the danger that it could be oxidized to a soluble form when the environment becomes oxygenated, use of the electrode enables the collection of insoluble U(IV) compounds and complete removal from the environment, followed by cleaning and re-use of the electrode (Gregory & Lovley, 2005).

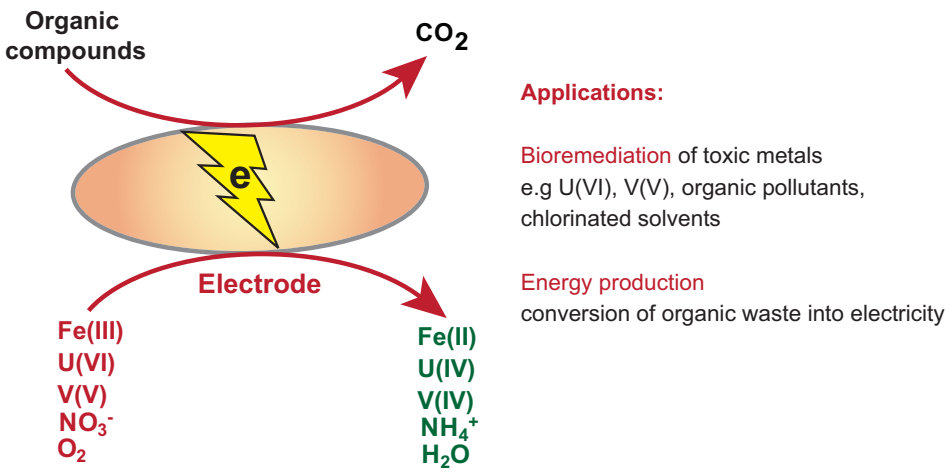


Figure 2. Physiological characteristics and potential applications of *Geobacter* species

Recently, it was discovered that in the process of extracting energy from organic compounds, which generates electrons, *Geobacter* species could transfer these electrons directly to the anodes of microbial fuel cells (MFCs), devices for current generation that employ bacteria as catalysts. Such MFCs produce a significant current, sufficient to power certain electronic devices (Bond *et al.*, 2002). This has led to the intensive study of *Geobacter* species to optimize the generation of electricity (Debabov, 2008). Studies focusing on understanding the electron transfer processes, so as to improve the yield of current, have provided important information. Based on available data, three major mechanisms have been proposed to be involved in electron transfer from cells to anode, including direct electron transfer via outer surface *c*-type cytochromes, long-range electron transfer via microbial nanowires, and electron flow through a conductive biofilm matrix containing cytochromes and soluble electron shuttles (Lovley, 2008a). Biofilm formation on the electrode appears to be important for efficient electricity generation by *Geobacter* species, and biofilm characteristics such as composition, adhesion, and conductivity of the biofilm are being studied. Identification of the factors that affect

biofilm formation could help us to understand the physiological properties of *Geobacter*, and therefore could enable the optimization of conditions for electricity production.

Chemotaxis genes have been previously implicated in various cellular processes, including regulation of cell motility (both pilus-based and flagellum-based), biosynthesis of the motility apparatus, and regulation of developmental genes. They could play important roles in *Geobacter* species for their survival in the environment as well as for their applications. We therefore focus our study on chemotaxis genes and their functions. The major aims of the study are:

1. To identify chemotaxis genes in *Geobacter* species. Based on this, to make educated guesses about their functions.
2. To identify the functions of chemotaxis genes in *Geobacter* species.

CHAPTER 2

IDENTIFICATION OF CHEMOTAXIS COMPONENTS IN THE GENOMES OF GEOBACTER SPECIES

Introduction

Chemotaxis is a trait shared by many bacteria that enables cells to move toward chemical attractants and away from repellents. The chemotaxis system of *E. coli* regulates flagellum-based motility; it has been studied in great detail and has served as a paradigm for chemotactic motility (Parkinson *et al.*, 2005, Falke & Hazelbauer, 2001). However, it is now apparent from genomic, genetic and biochemical studies conducted with other bacteria that a diversity of chemotactic signaling pathway functions and purposes exist well beyond the *E. coli* paradigm (Galperin, 2005, Szurmant & Ordal, 2004, Wadhams & Armitage, 2004).

There are 11 genes encoding proteins of the *E. coli* chemotaxis pathway, most of which are organized in a cluster near the flagellar genes (Blattner *et al.*, 1997). This cluster contains two of the five genes for the transmembrane chemoreceptors, which are also known as methyl-accepting chemotaxis proteins (MCPs), and a single gene for each of the chemotaxis signaling proteins: the autophosphorylation histidine kinase (CheA), a scaffold protein (CheW), a methyltransferase (CheR), a methylesterase (CheB), a response regulator (CheY), and CheY phosphatase (CheZ). The other three MCP genes are distantly located in the genome. Chemotactic signals are detected by a membrane array of MCPs, to which CheW and CheA are bound, and regulate CheA-mediated phosphorylation of CheY and CheB. By binding to the flagellar motor protein, FliM, CheY phosphate (CheY~P) induces swimming *E. coli* to tumble, which has the effect of

reorienting the direction of swimming. CheB~P reduces the kinase activity of CheA by demethylating the MCPs, which reduces the rate of CheY~P (and CheB~P) formation, and consequently reduces the cell tumbling frequency. The tumble-promoting activity of CheY~P is also extinguished by the action of CheZ. Overall, this stimulus-response pathway biases swimming motion of *E. coli* toward attractants and away from repellents. Adaptation to stimuli, mediated by the reversible methylation of MCPs in the process catalyzed by CheR and CheB, allows cells to remain sensitive to small changes in chemoeffector concentration over a large range (Antommattei & Weis, 2006, Falke *et al.*, 1997).

Geobacter species are δ-*Proteobacteria* that are often the predominant species in a variety of sedimentary environments where Fe(III) reduction is important. Their ability to remediate contaminated environments and to produce electricity makes them attractive for further study. Cell motility, biofilm formation, and type IV pili all appear important for the growth of *Geobacter* in changing environments and for electricity production. Recent studies in other bacteria have demonstrated that signaling pathways homologous to the paradigm established for *Escherichia coli* chemotaxis can regulate type IV pilus-dependent motility, the synthesis of flagella and type IV pili, the production of extracellular matrix material, biofilm formation, and the regulation of developmental genes. The classification of these pathways by comparative genomics improves one's ability to understand how *Geobacter* thrives in natural environments and how to improve the use of *Geobacter* in microbial fuel cells.

Methods

For protein sequence similarity searches, NCBI protein BLAST and position-specific-iterated-BLAST (blastp and psi-blast, respectively, http://blast.ncbi.nlm.nih.gov/) (Altschul *et al.*, 1997) were used with default parameter values against the genomes of *G. sulfurreducens* PCA, *G. metallireducens* GS-15, and *G. uraniireducens* Rf4 (GenBank accession numbers AE017180.1, CP000148.1 and CP000698.1, respectively, www.ncbi.nlm.nih.gov/genomes/lproks.cgi). To identify the *Geobacter* homologs of chemotaxis genes, *E. coli, Bacillus subtilis* and *Thermotoga maritima* chemotaxis proteins were used as the test sequences, because these proteins are well-studied representatives, and are listed in the curated databases (Letunic *et al.*, 2002, Marchler-Bauer *et al.*, 2005, Finn *et al.*, 2006). The following sequences were used: the *E. coli* aspartate receptor methyl-accepting (MA) domain (residues 267-514) (gi|16129838), the complete sequences of *E. coli* CheA (gi|1788197), CheB (gi|16129835), CheR (gi|16129836), CheW (gi|16129839), CheY (gi|16129834) and CheZ (gi|16129833); the complete sequences of *B. subtilis* CheC (gi|2634017), CheD (gi|2634018) and CheV (gi|2633772), and the complete sequence of *T. maritima* CheX (gi|81553634). ClustalW (www.ebi.ac.uk/clustalw/) was used with default values for the parameters to conduct multiple sequence alignments to determine percent identities and to establish the class membership of the methyl-accepting domains (Thompson *et al.*, 1994). TMHMM2 (Krogh *et al.*, 2001), TmPred (Stoffel, 1993), and TopPred (Claros, 1994) were used (with parameters set to default values) to predict the number of transmembrane helices in the putative methyl-accepting chemotaxis proteins. A polypeptide segment was designated a transmembrane α-helix when at least two of the

three programs identified the same polypeptide segment as a transmembrane helix. Phylip (version 3.6) was used to construct CheA and CheY phylogenetic trees by the neighbor-joining method (Felsenstein, Felsenstein, 1989), as implemented in NEIGHBOR. SEQBOOT was used to generate 1000 bootstrap replicates and pairwise distances were estimated with PROTDIST. The JTT model was used with no among-site variation. The trees were left unrooted.

The organization of *che* gene operons in *Geobacter* species was predicted with FGENESB (Softberry Inc., www.softberry.com). FGENESB identifies protein-coding genes with Markov chain models of coding regions and translation start and termination sites, and annotates them with information from public databases. The sequence parameters (coding content, oligonucleotide composition, and gene length distribution) were estimated in FGENESB for each genome separately through an iterative procedure with the minimum ORF length set to 100 nt. Additional features, *e.g.* tRNA and rRNA, σ^{70} family promoters, and Rho-independent terminators, were predicted from sequence similarity, linear discriminant analysis, or modeling approaches. FGENESB-based operon predictions were generated from the directions of adjacent genes, the distribution of intergenic distances, the presence or absence of predicted promoter and terminator regions, and the conservation of pairs of adjacent genes across microbial genomes (V. Solovyev, *personal communication*). The operon annotation of the *G. sulfurreducens* genome used in this study has been described previously (Krushkal *et al.*, 2007), and is available online (www.geobacter.org/research/gsel/) (Krushkal *et al.*, 2008).

σ^{54}-regulated promoters were predicted from a search of the *G. sulfurreducens* genome with PromScan (Studholme *et al.*, 2000). This software assigns a score

representing the Kullback-Leibler distance, based on 186 known sites from 47 bacterial species (Barrios *et al.*, 1999). The *G. sulfurreducens* genome was found to contain 110 predicted σ^{54}-regulated promoters with a score equal to or greater than 80 (the default value) in noncoding regions upstream of genes and operons. The current accuracy of prediction is 78%, an estimate obtained from experiments that positively identified 14 RpoN-dependent regulation sites out of 18 predicted sites (J. Krushkal, C. Leang, M. Puljic, T. Ueki, R. Adkins, and D. Lovley, *unpublished results*). In addition, PromScan was used to look for σ^{54}-regulated promoters upstream of the major *che* clusters in the *G. metallireducens* and *G. uraniireducens* genomes. Finally, putative σ^{28}-regulated promoters upstream of the flagellar filament gene (*fliC*) and the major *che* clusters in the genomes of *G. sulfurreducens*, *G. metallireducens* and *G. uraniireducens* were identified with Virtual Footprint (Munch *et al.*, 2005) and the Neural Network Promoter Prediction software for bacterial species (Reese *et al.*, 1996). Five hundred base pairs upstream of the putative initiation codons of genes of interest were analyzed using default parameters.

Results and Discussion

Geobacter Chemotaxis Genes: Numbers and Organization

BLAST analysis of the *G. sulfurreducens*, *G. metallireducens*, and *G. uraniireducens* genomes identified multiple copies of the chemotaxis genes; over 60 genes in each species were homologous to the known *che* and *mcp* genes in *E. coli*, *B. subtilis* and *T. maritima* (Table 1). Homologs of all the *che* genes from *E. coli* were present in the *Geobacter* species, except *cheZ*, which is found much more frequently in genomes of β- and γ-*proteobacteria* in comparison to the genomes of α-, ε-, and especially δ-*proteobacteria* (Wuichet *et al.*, 2007). The *Geobacter* genomes also contained *cheC*, *cheD*, *cheV*, and *cheX* homologs. With the exception of the genes for the chemoreceptors – the methyl-accepting chemotaxis proteins (MCPs), which were dispersed throughout the genomes, most of the *che* genes were clustered, as shown in Figure 3. In some cases, additional genes encoding hypothetical proteins of unknown function or annotated proteins with functions not known to be involved in chemotaxis-related signaling pathways were located in these clusters. There are six major chemotaxis-related gene clusters in *G. sulfurreducens*, and seven major clusters each in *G. metallireducens* and *G. uraniireducens*; their physical arrangements are depicted in Figure 3. None of these clusters is located close to the flagellar gene clusters.

Gene	Species				
	E. coli	***B. subtilis***	***G. metallireducens***	***G. sulfurreducens***	***G. uraniireducens***
cheA[b]	1	1	5	4	7
cheB	1	1	8	4	5
cheR	1	1	9	5	10
cheW	1	1	8	10	10
cheY[c]	1 (1)	1 (3)	10 (21)	7 (25)	10 (25)
cheZ	1	0	0	0	0
cheC	0	1	2	1	1
cheD	0	1	3	3	2
cheX	0	0	1	1	1
cheV	0	1	1	1	1
mcp	5	10	18	34	24
Total	11	17	65	70	71
No. of *che* clusters[d]	1	1	7	6	7

Table 1. Numbers of *che* gene homologs in *E. coli*, *B. subtilis* and *Geobacter* species.[a]

[a]homologs numbers were determined by blastp searches (with default values for the parameters).

[b]The numbers for *cheA* in the genomes of *G. metallireducens* and *G. uraniireducens* each include a contribution from one *cheAY* fusion.

[c]The numbers are *cheY* genes in the major clusters. Numbers in parentheses also includes genes that encode singleton CheY-like receiver domain proteins.

[d]Chemotaxis gene clusters are defined to contain three or more *che* genes.

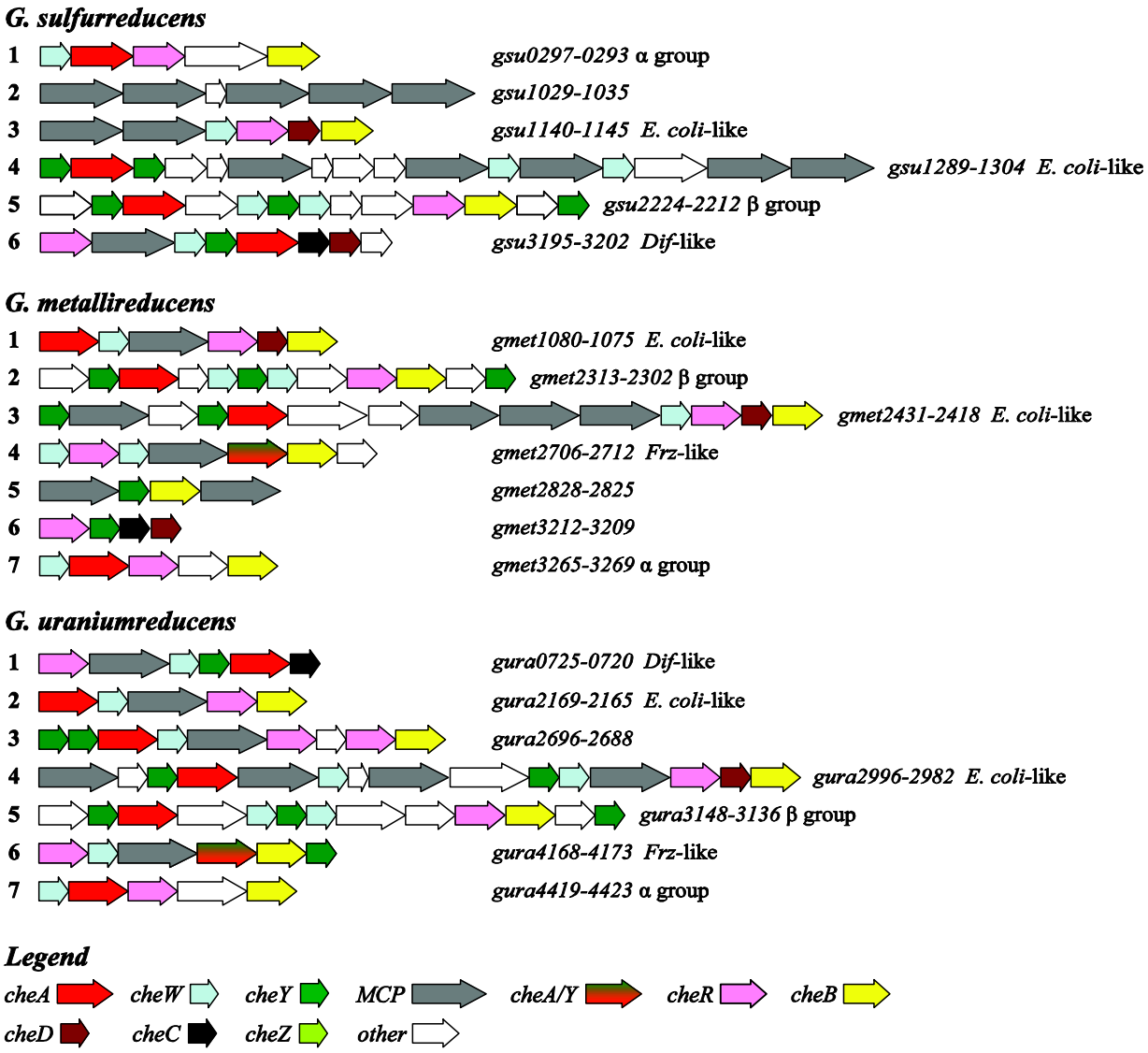


Figure 3. Physical arrangement of the major *Geobacter* chemotaxis-like gene clusters. Affiliations with *che* clusters of known function are indicated after the clusters, as *E. coli*-like, *Dif*-like and *Frz*-like (both from *M. xanthus*), and the α and β *group*s. These assignments were made by the relative agreements between *che* gene content, the physical arrangement in the cluster and the percent identities. The α and β group designations refer to *che* clusters that are unique to the *Geobacteraceae* and the *δ-Proteobacteria*, respectively.

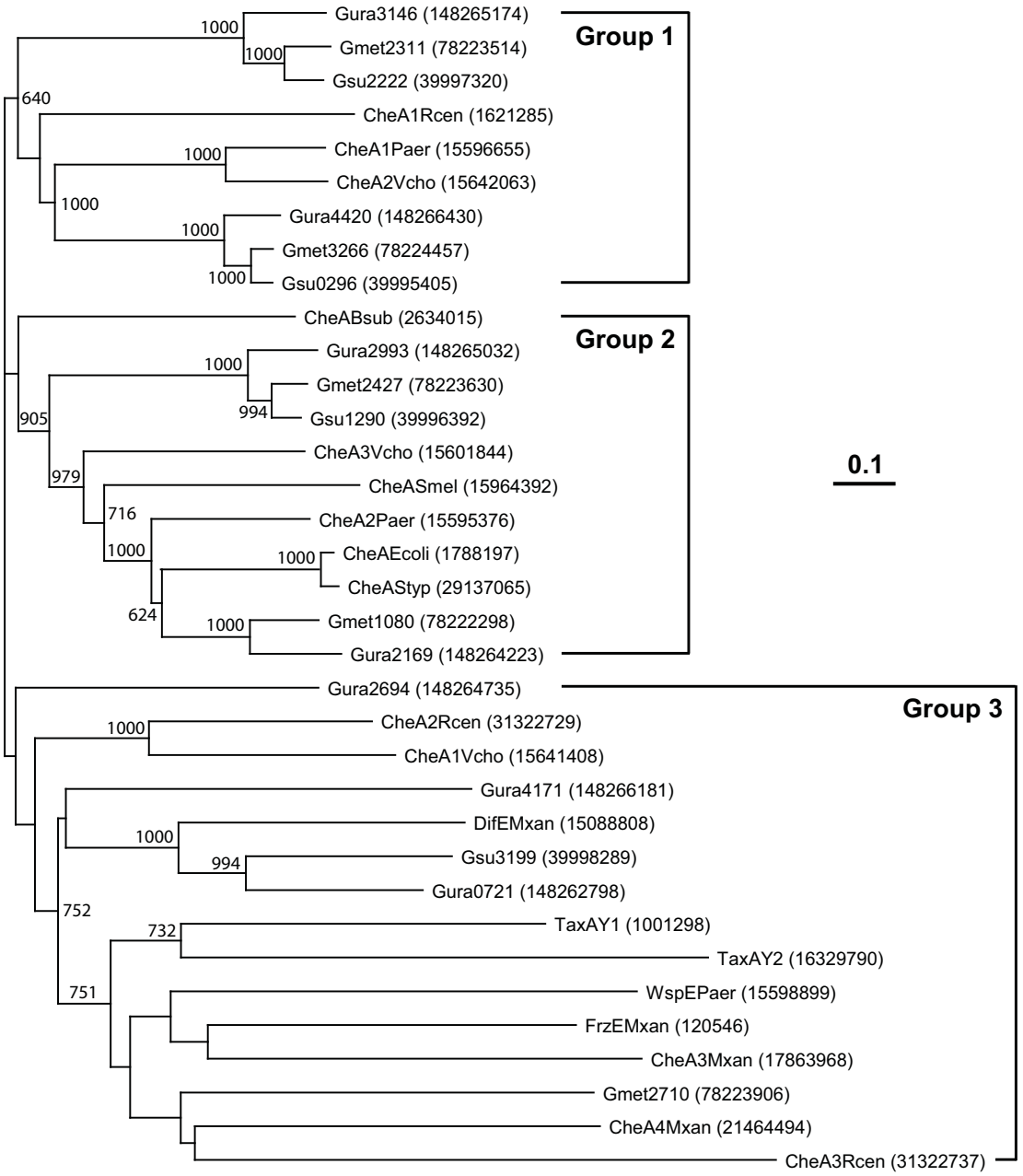


Figure 4. Neighbor-joining tree of putative CheA homologs of *Geobacter* species and CheAs from other well-studied species. These include *E. coli* (Ecoli), *B. subtilis* (Bsub), *P.. aeruginosa* (Paer), *S. meliloti* (Smel), *R. centenum* (Rcen), *V. cholerae* (Vcho), *M. xanthus* (Mxan), *Salmonella typhimurium* (Styp) and *Synechocystis* sp. strain PCC6803 (Syne). The GenInfo Identifier protein sequence numbers are displayed in parentheses at right. All positions with gaps in the aligned sequences were excluded. Bootstrap values from 1000 replicates of >600 are shown in respective nodes. The tree figure was generated with TreeView, version 1.6.6 (Page, 1996).

The genomes of *G. sulfurreducens*, *G. metallireducens* and *G. uraniireducens* encode four, five and seven predicted *cheA* genes, respectively. The homologs enoded by the *cheA* genes are clustered in three groups of the phylogenetic tree (Figure 4), demonstrating that the multiple *cheA* genes did not result from recent gene duplication events, but are paralogs that have been evolving separately for some time, which suggests that they play distinct cellular roles. Each CheA homolog, together with the other cognate *che* gene products, is likely to regulate a separate chemotaxis-like pathway. The presence of multiple *che* homologs and clusters is a strong indication of different pathways that raises intriguing questions about function, and whether or not the pathways are redundant or exhibit crosstalk. By comparing the gene order and the percent identities of gene products with other bacteria, in which chemotaxis and chemotaxis-like signaling pathways have been studied extensively, we were able to predict the functions for many of the *Geobacter che* clusters. From this analysis, it seems unlikely that different clusters constitute redundant pathways; instead, each pathway has a distinct function. In addition, plausible mechanisms to reduce unwanted crosstalk between pathways emerged.

The *Geobacter* genomes are predicted to have large numbers of standalone response regulators proteins that are comprised only of the receiver domain (Galperin, 2006); we refer to these as CheY-like proteins. The *G. metallireducens*, *G. sulfurreducens*, and *G. uraniireducens* genomes have 21, 25, and 25 homologs, respectively, but the majority are probably *not* involved in chemotaxis-like signaling (Wuichet et al., 2007). Only 38% of the homologs are located in the major *che* or flagellar gene clusters (Table 1), the remainder (11, 18 and 15, respectively) are located elsewhere on the chromosome. Of those we suspect to play a role in chemotaxis-like

signaling, *i.e.* the *cheY* genes that are located in the major *che* or flagellar gene clusters, about 50% reside in a branch of the phylogenetic tree with *E. coli* and *Salmonella* CheY (four apiece from *G. metallireducens* and *G. sulfurreducens*; five from *G. uraniireducens*, Figure 5). These CheY homologs are most likely to have response regulator functions as the substrates of CheA-mediated phosphorylation in chemotaxis pathways. The *Geobacter* CheY-homologs that are located elsewhere in the tree (relatively distant to *E. coli* and *Salmonella* CheY), but are situated in *che* or flagellar gene clusters on the chromosome, probably also function in chemotaxis-like pathways, perhaps in some other manner. By contrast, the genes encoding the most distantly related CheY-like proteins, *i.e.* located outside *che* clusters, away from the flagellar genes, *and* (relatively) distant to *E. coli* and *Salmonella* CheY in the phylogenetic tree (Figure 5), probably function in other two-component pathways. For instance, *B. subtilis* and *Nostoc* species CheY-like homologs, which are not in the *che* clusters, are involved in two-component pathways unrelated to chemotaxis (Campbell *et al.*, 1996, Fabret *et al.*, 1999). Therefore, we postulate that only the standalone receiver proteins encoded within the *Geobacter che* and flagellar gene clusters plausibly represent CheYs with functions in chemotaxis-like signaling pathways.

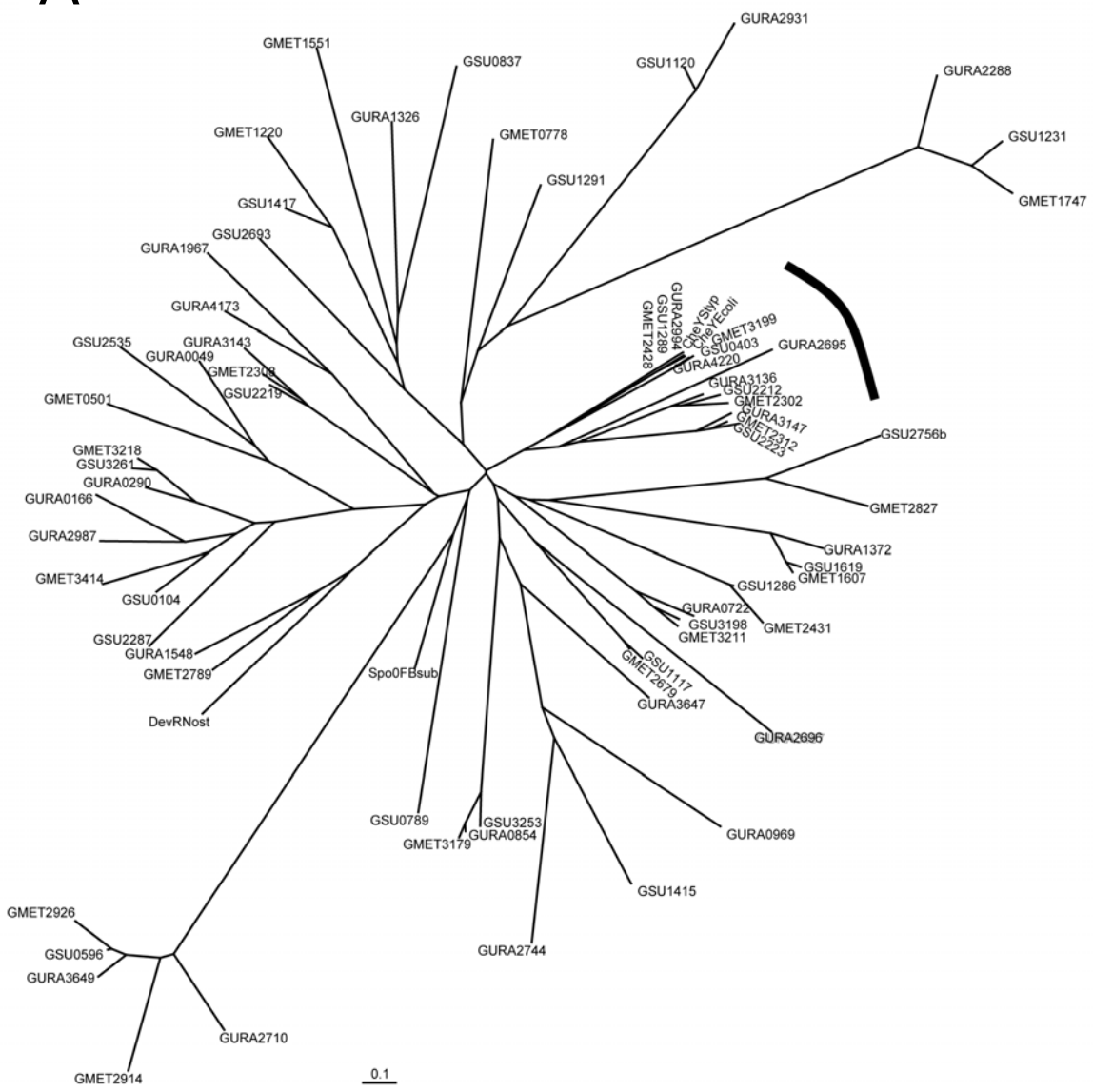


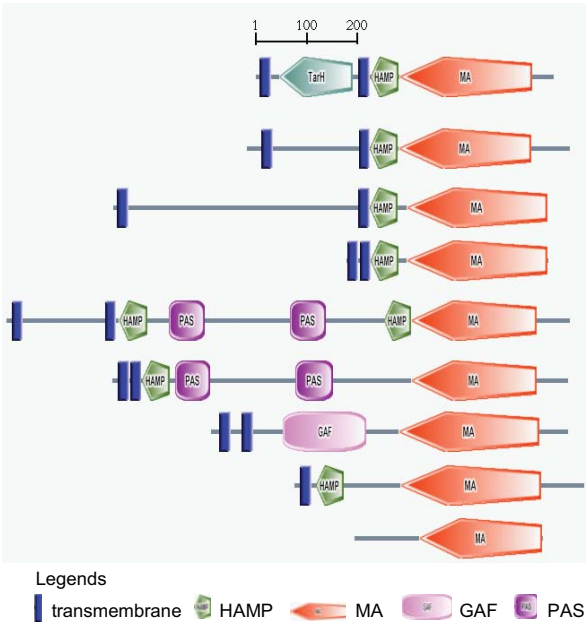
Figure 5. A neighbor-joining phylogenetic tree of the *Geobacter* CheY homologs and selected CheY & CheY-like proteins from other species: *E. coli* (Ecoli), *B. subtilis* (Bsub), *Salmonella enterica* subsp. enterica serovar Typhi Ty2 (Styp), and *Nostoc* sp. strain ATCC 29133 (Nost)

B	*Geobacter* CheY Homologs			
	***cheY*s in *che* and *fla* clusters**		***cheY*s outside *che* and *fla* clusters**	
Species	***E. coli* Branch**	**Other Branches**	***E. coli* Branch**	**Other Branches**
G. sul	4/7	3/7	0/15	18/18
G. met	4/10	6/10	0/15	11/11
G. ura	5/10	5/10	0/15	15/15

Table 2. *Geobacter* CheY homolog distributions between the *E. coli* CheY branch and the rest of the phylogenetic tree. Approximately 50% of the homologs situated on the chromosome within the major *che* or flagellar gene clusters reside in a tree branch with the *E. coli* and *S. typhimurium* CheYs (marked by a black arc). No CheY homologs situated on the chromosome outside *che* or flagellar gene clusters reside in the same branch with *E. coli* CheY.

Number and Diversity of *Geobacter MCPs*

The three *Geobacter* genomes investigated in this study were found to have significant numbers of genes for MCPs: 34 in *G. sulfurreducens*, 18 in *G. metallireducens*, and 24 in *G. uraniireducens* (Table 1). These putative MCPs were identified through the presence of the highly conserved methyl-accepting (MA) domain, which was first assigned a biochemical function in the *E. coli* chemoreceptors as the domain methylated in a CheR-dependent process (Kort *et al.*, 1975, LeMoual & Koshland, 1996). The large number of MCP-coding genes in the *Geobacter* genomes, by comparison to either *E. coli* or *B. subtilis*, plausibly reflects a greater need to detect sensory stimuli in the subsurface environment. With the exception of the aerotaxis receptor, all *E. coli* MCPs have periplasmic ligand-binding domains that detect the external chemoeffector concentrations, two transmembrane (TM) helices, and the (juxtamembrane) HAMP and methyl-accepting (MA) domains located in the cytoplasm. The sequences of the predicted *Geobacter* MCPs reveal significantly greater diversity in the domain organization and architecture of the sensing domains (Figure 6).



G. met (18)	*G. sul* (34)	*G.ura* (24)	
			E.Coli Tar
11	25	15	2TMs, regular periplsam (~150˜200 aa)
1	1	1	2TMs, large periplasm (~350-450 aa)
2	3	3 (+1)	2TMs, small periplasm (<20 – (~100 aa))
1	0	1	2TM, regular periplasm and PAS domain
0	0	1	Aer like
1	1	0	2TMs, GAF domain
0	1	0	1 TM
2	3	2	Soluble

Figure 6. Domain architectures of MCPs in *Geobacter* sp. show diversity. *Geobacter* MCPs are classified according to the periplasmic domain size, the length of the methyl-accepting (MA) domain (LeMoual *et al.*, 1997) and the presence of the PAS (Ponting & Aravind, 1997) (Gmet_2422, Gura_1191, Gura_2989) or the GAF domain (Aravind & Ponting, 1997) (GSU1704, Gmet_1641).

Number of Transmembrane Helices	**Size of Periplasmic Domain** (amino acid residues)	**MCP of Each Type (%)**		
		G. sul	***G. met***	***G. ura***
2	Large (250 to 430)	3	0	4
	Medium (150 to 200)	**76**	**72**	**67**
	Small (< 100)	9	17	21
1		3	0	0
0 (Soluble)		9	11	8

Table 3. Distribution of *Geobacter* MCPs by number of transmembrane segments & size of periplasmic domains.

The N-terminal regions of MCPs sense various environmental stimuli through diverse means; the length and heterogeneity of these regions are greater compared to the cytoplasmic domains, which are mostly organized like the *E. coli* MCPs (a single HAMP domain followed by the MA domain). Domain architectures of representative *Geobacter* MCPs are shown in Figure 6 and Table 3. With respect to transmembrane (TM) segments, the *Geobacter* MCPs fall into three groups according to the number of predicted TM helices (zero, one or two). Ninety percent of the *Geobacter* MCPs have two TM helices. Most of these have periplasmic domains that are ~150-200 amino acid residues (aa) in length (80 %), which are most similar to the periplasmic domains of the major *E. coli* MCPs. Three percent of the *Geobacter* MCPs have larger periplasmic domains (~250-430 aa), while the others have a significantly smaller domain (< 100 aa). MCPs with the Tar-like and larger periplasmic domains probably detect signals through these domains by ligand binding. On the other hand, the MCPs with small periplasmic domains are more likely to detect signals via associations with other proteins, as in the case of DifA of *M. xanthus* (Black *et al.*, 2006, Bonner et al., 2005), and MCPs that have no TM segments are likely to detect intracellular signals (Wadhams & Armitage, 2004). MCP MA domains were recognized to belong to a superfamily based on a multiple sequence alignment first conducted by LeMoual and Koshland (LeMoual & Koshland, 1996). A more recent analysis of approximately 2000 MCPs identified seven classes (named 24H, 28H, 34H, 36H, 38H, 40H and 44H), which are defined by the number of heptad repeats (H) in the cytoplasmic domain (Alexander & Zhulin, 2007). The most well-characterized MCPs of *E. coli*, Tar and Tsr, belong to class 36H, and the MCPs from *T. maritima* (TM1143) and *B. subtilis* (McpA and McpC) belong to class 44H.

Multiple sequence alignments of the *Geobacter* MCPs revealed that *G. metallireducens* has MCPs in classes 24H, 34H, 36H and 40H; *G. sulfurreducens* has MCPs in classes 24H, 34H, 40H and 44H, and *G. uraniireducens* has MCPs in classes 24H, 34H, 36H, 40H and 44H. The majority of the MCPs are members of class 34H (17, 24, and 21 % in *G. metallireducens*, *G. sulfurreducens* and *G. uraniireducens*, respectively) or class 40H (61, 71 and 46 %, respectively, see alignments in Appendix A). *G. metallireducens* and *G. uraniireducens* each have one MCP in class 36H (Gmet_1078, Gura_2167), and *G. sulfurreducens* and *G. uraniireducens* each have one MCP in class 44H (GSU3196, Gura_0724). MCPs and Che proteins form specific clusters. In *E. coli,* all the MCPs and most of the Che proteins are found in clusters that are often located at the cell poles (Cantwell *et al.*, 2003, Maddock & Shapiro, 1993b, Sourjik & Berg, 2000). When bacteria have two or more chemotaxis (or chemotaxis-like) gene clusters, the signaling protein clusters are observed to have distinct locations and compositions (Guvener & Harwood, 2007, Guvener *et al.*, 2006, Wadhams *et al.*, 2003). We speculate that MCP class membership is a contributing factor of cluster specificity. According to this reasoning, MCPs in the same class are more likely to belong to the same cluster, and conversely, MCPs in different classes are likely to segregate. Cluster formation, in part, is generated by contacts between MCP MA domains, each consisting of a coiled-coil hairpin that dimerizes to form a long four-helix bundle (Kim *et al.*, 1999, Park *et al.*, 2006), the bundle length being determined by the number of heptad repeats: ~210 Å for class 36H MCPs (*E. coli* Tsr) and ~260 Å for class 44H MCPs (*T. maritima* TM1143). We postulate that class-specific MCP heterodimers are more likely to form for the following reason: two different MCPs, which contain MA domains belonging to the same

class, are more likely to engage in the interactions that lead to the formation of a four-helix bundle than two MCPs that contain MA domains from different classes (i.e., having different MA domain lengths).

The localization of *P. aeruginosa* and *R. sphaeroides* chemotaxis protein clusters provides some possible examples of class-specific MCP segregation. *P. aeruginosa* McpB and WspA, which are found in distinct signaling complexes in distinct locations (polar and lateral, respectively), belong to different classes (36H and 40H, respectively) (Guvener & Harwood, 2007, Guvener et al., 2006). *R. sphaeroides* McpG and TlpT (a soluble MCP) belong to different MA classes (34H and 36H, respectively) and locate in different clusters (polar membrane and cytoplasmic locations, respectively) (Wadhams et al., 2003). We anticipate that the multiple classes of MCPs present in *Geobacter* species contribute to the formation of segregated MCP signaling clusters. On the other hand, MA class membership is certainly not the only factor to consider. For example, this mechanism cannot easily explain the localization of MCPs that do not belong to any class (Alexander & Zhulin, 2007). In addition, the compositions of signaling clusters are influenced undoubtedly by the specificity of interactions between the different MA domain and Che protein homologs. These effects (and others), considered together, can contribute to the assembly of specific signaling units, which function in the same cell without unwanted crosstalk.

The Prevalence and Specificity of CheR Tethering Segments

A semi-conserved pentapeptide (NWETF) at the C-terminus of some MCPs, first observed in the *E. coli* high abundance receptors Tar and Tsr (Wu *et al.*, 1996), has a well-established role in sensory adaptation by mediating efficient receptor methylation

and demethylation (Antommattei & Weis, 2006, Barnakov *et al.*, 2001, LeMoual et al., 1997, Li *et al.*, 1997, Wu et al., 1996). In the process of receptor methylation, the pentapeptide NWETF binds to the β-subdomain of CheR at a location that is distinct from the active site–methylation site interaction that tethers CheR near the methylation sites in clustered receptors (Djordjevic & Stock, 1998). It is plausible that all MCPs containing the C-terminal NWETF or a pentapeptide similar to NWETF provide adaptational assistance via the mechanism established in *E. coli* (Antommattei *et al.*, 2004, Li et al., 1997). MCPs that contain the CheR-binding pentapeptide are restricted primarily to the α-*Proteobacteria*; the genomes of bacteria in other phyla reveal few, if any, MCPs that contain a recognizable CheR-tethering segment, as defined previously (Antommattei & Weis, 2006). In such species – for example *B. subtilis* and *T. maritima*, methylation operates through a different, pentapeptide-independent mechanism (Perez & Stock, 2007). Less than 10% of the ~2500 MCPs listed in the SMART database of completed bacterial genomes contain a recognizable CheR tethering segment; this segment always follows the MA domain (SM00283) in the primary sequence of the MCP, which then ends in a pentapeptide that binds CheR (Antommattei et al., 2004). Therefore, many MCPs are probably methylated and demethylated *via* a pentapeptide-independent mechanism.

Closer analysis of all the MCPs that contain the NWETF pentapeptide or a similar pentapeptide reveals a restricted class membership of either class 34H or class 36H (Alexander & Zhulin, 2007). Eighty-five percent of these MCPs belong to class 36H *and* contained the class-specific xWxxF pentapeptide motif; 15% belong to class 34H *and* contained the class-specific xF/YxxF/Y motif for the pentatpeptide (Antommattei &

Weis, 2006). In contrast to the kingdom-wide percentages, most pentapeptide-containing *Geobacter* MCPs belonged to class 34H (100%, 75%, and 80% for *G. sulfurreducens*, *G. metallireducens* and *G. uraniireducens*, respectively). *G. metallireducens* and *G. uraniireducens* have one *mcp* gene apiece in the class 36H with a C-terminal DWKEF pentapeptide, a sequence more similar to the *E. coli* consensus (NWETF). Using the pentapeptide-containing MCPs as one criterion, we designated the *che* clusters to which these *mcp* genes belong as '*E. coli*-like' clusters (Figure 3).

To identify possible class-specific MCP-methyltransferase tethering interactions, we compared the aligned β-subdomain sequences of the *Geobacter* CheR homologs to the *Salmonella* and *E. coli* CheR sequences. The *Salmonella* CheR structure, co-crystallized with the NWETF pentapeptide, has enabled the identification of residues in the β-subdomain that are involved in the peptide-CheR interaction (Q182, G188, R187, G190, G194 and R197, numbered according to *Salmonella* CheR, PDB# 1bc5) (Djordjevic & Stock, 1998, Perez & Stock, 2007). Figure 7 shows aligned sequences from the β-subdomains of all the *Geobacter* CheR homologs, together with the *E. coli* and *Salmonella* sequences (residues 166-199). Using this alignment, we divided the *Geobacter* CheRs into three groups. Two groups (A and B) displayed significant identity of the residues important for binding a pentapeptide; the third and largest group (C) did not (Figure 7). Consequently, we concluded that the CheR homologs in Group C probably do not methylate MCPs by the *E. coli* mechanism.

The colocalization of *mcp* and *cheR* genes within the same clusters was consistent with the hypothesis that CheR homologs in groups A and B bind to MCPs containing a C-terminal pentapeptide; these CheR homologs are located in *che* gene clusters

containing at least one gene that encodes a pentapeptide-containing MCP (Figure 3). The two CheR homologs that comprise group A are located adjacent to class 36H MCPs (Gmet_1078, Gura_2167), and have DWKEF as the C-terminal pentapeptide – judged to be more similar to the *E. coli* consensus (NWETF). By contrast, the consensus pentapeptide coded by *mcp* genes located in the *che* gene clusters with group B CheR homologs is EFEKF. All 14 MCPs that contain this consensus pentapeptide belong to class 34H (Figure 7), and 10 of these are located in the *che* gene clusters that contain the group B *cheR* genes (Figure 3). Differences in the consensus pentapeptide for class 34 *versus* class 36 MCPs correlate with differences in the β-subdomain amino acid residues in pentapeptide-binding pockets of group A *versus* B CheR homologs (respectively, Figure 7). Thus, it is plausible that these differences contribute to (and reflect) class-specific MCP-CheR interactions.

By contrast, all group C CheR homologs are either (a) not located in a *che* gene cluster, (b) located in *che* clusters that do not contain an *mcp* gene, or (c) located in *che* clusters that contain genes encoding class 40H or 44H MCPs. These MCPs do not contain recognizable CheR tethering segments (terminating in a pentapeptide) according to criteria defined previously (Antommattei & Weis, 2006). Thus, it is probable that group C CheRs use a pentapeptide-independent mechanism for receptor methylation, similar to that observed with *T. maritima* (Perez & Stock, 2007). In addition, we interpret the specific pairings within the *che* gene clusters, of the CheR groups (A, B, C) and the MCP classes (36H, 34H, 40H/44H, respectively), as support for the idea of class-specific receptor signaling.

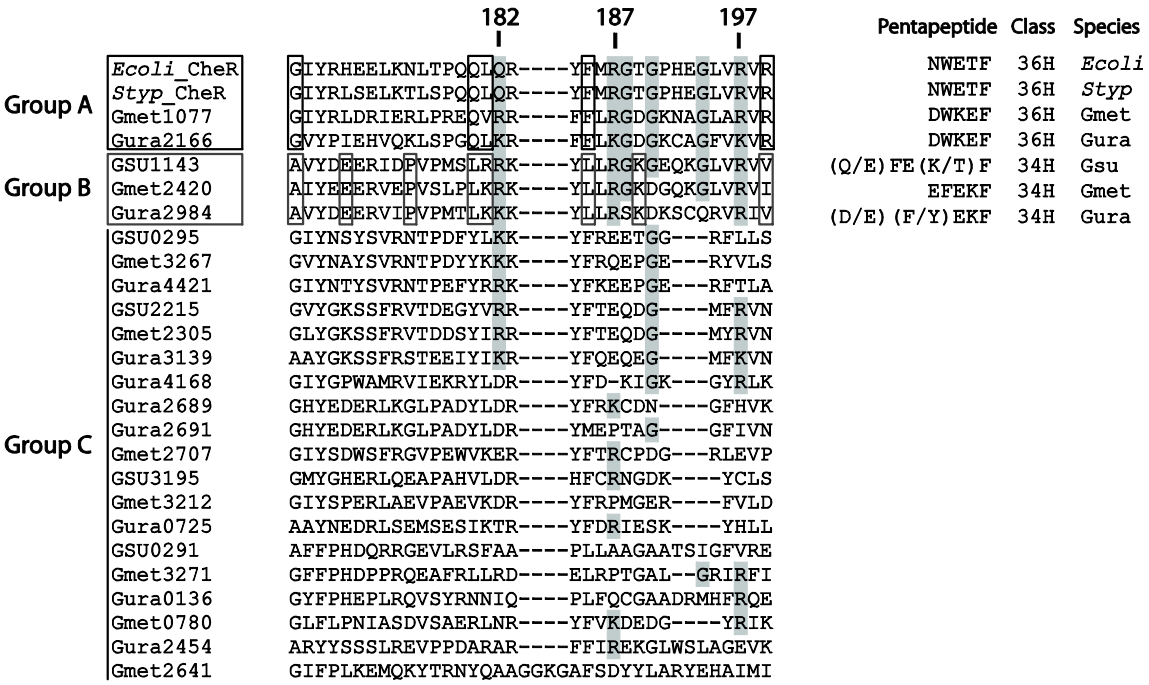


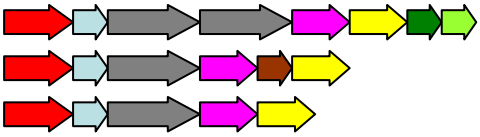
Figure 7. Alignment of the beta-subdomain of *Geobacter* CheR homologs with *E. coli* and *S. enterica* CheR. Based on sequence alignments, the *Geobacter* CheRs were divided into three groups. Two groups (A and B) displayed significant identity of residues important for pentapeptide binding (highlighted in grey) and the third group (C) did not. Gene positioning provides further evidence that the group A and B homologs bind to MCPs containing the C-terminal pentapeptide: these homologs are located in *che* clusters with pentapeptide-containing MCPs (Figure 3). Group A consists of two CheR homologs that are located near two class 36H MCPs. The consensus pentapeptide of the MCPs that are cognate to the Group B CheR homologs, EFEKF, is found in class 34H MCPs.

Predicted Function of *E. coli*-like Chemotaxis Clusters

Above, we designated *Geobacter che* clusters operationally as '*E. coli*-like' by the presence of one or more *mcp* genes that encode MCPs with CheR tethering segments. In addition, the *Geobacter* CheA homologs in these clusters belong to the same phylogenetic grouping as *E. coli* CheA (Figure 4). These clusters were sorted further into two types; type 1 clusters – clusters one and two in the *G. metallireducens* and *G. uraniireducens* genomes, respectively (Figure 3), have significant resemblances to the *E. coli mocha-meche* cluster, judged by the gene order and by the percent identities between predicted *Geobacter* proteins and the *E. coli* proteins (Figure 4 and Figure 8). Notably, the *Geobacter mcp* genes in these two clusters encode MCPs that belong to class 36H, the same as *E. coli* MCPs.

Type 2 clusters are also characterized by significant sequence identity (although lower than Type 1), but the gene positions bear a comparatively small resemblance to the *E. coli* cluster. Moreover, the Type 2 clusters contain predicted ORFs in significant numbers that have unknown function or assigned functions other than chemotaxis (See Figure 8 for comparisons of gene arrangement and the percent identities of individual gene products). The Type 2 clusters possess multiple genes coding for MCPs that belong to class 34H; many of these contain a CheR-tethering segment that terminates in an 'NWETF-like' pentapeptide at the C-terminus.

A. Type 1 gene arrangements



B. Percent identities of the *Geobacter* Type 1 cluster proteins with *E. coli*. The MCP identities were determined between the *E. coli* aspartate receptor cytoplasmic domain (c-Tar) and the corresponding c-domains of *Geobacter* MCPs. For CheA, CheW, CheR and CheB the full-length gene products were used.

E. coli	**CheA**	**CheW**	**Tar Cyt. Domain (c-Tar)**	**CheR**	**CheB**
G. met (cluster 1)	47	63	53	50	62
G. ura (cluster 2)	46	60	53	48	60
Mean identities with remaining *Geobacter* homologs[a]					
G. met	32 ± 5 (4)	27 ± 5 (8)	29 ± 6 (7)	29 ± 5 (5)	37 ± 8 (5)
G. ura	32 ± 3 (6)	27 ± 4 (6)	27 ± 6 (7)	32 ± 4 (7)	41 ± 7 (5)

[a]mean ± standard deviation (number of homologs in the average); including Type 2 clusters.

C. Type 2 gene arrangements

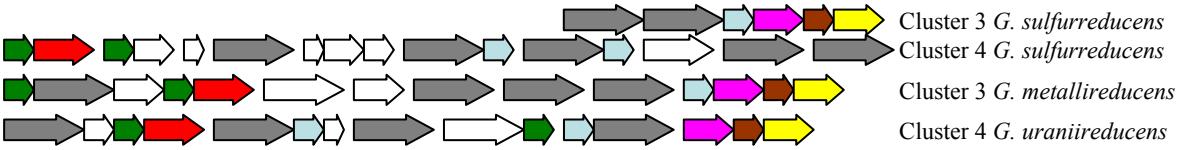


Figure 8. Gene arrangement of *E. coli*-like clusters in *Geobacter* (A, C), with the percent identity of the individual gene products to *E. coli* (B, D).

D. Percent identities determined between the gene products of Type 2 clusters and *E. coli* proteins (as described in Figure 8B). For CheW and CheY, the percent identities of both loci are reported. For the MCPs, the mean and standard deviation of the identities are reported, and the number of sequences used is in parentheses.

E. coli	**CheA**	**CheW**	**c-Tar**[a]	**CheR**	**CheB**	**CheY**
G. sul *(clusters 3 & 4)*	37	33 ± 1.5 (3)	35 ± 3 (7)	41	50	26, 37
G. met *(cluster 3)*	36	31	33 ± 1 (3) 15 (1)	38	50	23, 37
G. ura *(cluster 4)*	35	29, 30	32 ± 1 (3) 14 (1)	37	50	25, 35
Mean identities with remaining *Geobacter* homologs[b]						
G.sul	32 ± 2 (3)	27 ± 7 (6)	25 (1)	29 ± 5 (3)	38 (2)	29 ± 3 (4)
G. met	30 ± 6 (3)	27 ± 4 (5)	29 ± 1 (3)	26 ± 7 (4)	35 ± 3 (4)	29 ± 5 (5)
G. ura	32 ± 4 (5)	26 ± 5 (6)	27 ± 2 (3)	27 ± 6 (6)	37 ± 3 (4)	26 ± 7 (6)

[a]mean ± standard deviation (number of homologs in average). One c-domain of lower identity, as noted, was excluded from the average.
[b]mean ± standard deviation (number of homologs in average); excluding Type 1 clusters.

Three features distinguish the *E. coli*-like *Geobacter che* clusters from the *E. coli* cluster. (*1*) The *E. coli meche* operon contains *cheZ*, but *Geobacter* genomes do not, and so dephosphorylation of CheY in *Geobacter* must be CheZ-independent. (*2*) Multiple CheW genes are found in each cluster (except for *G. metallireducens*), an observation made previously with other bacteria. Studies of the CheW homologs in *R. sphaeroides* have led to the suggestion that these homologs do not perform redundant functions, but engender MCP-specific interactions, a proposal based on observed differences in binding affinity (Martin et al., 2001). It has also been suggested that multiple CheWs allow additional MCPs to be incorporated within the chemosensory system, since there tend to be more *mcp* genes than *cheW* genes (Whitchurch et al., 2004). Another interesting hypothesis has been proposed: the different CheWs may recognize MCPs in a class-specific manner, producing different specific signaling pathways in a Che protein

complex (Wuichet et al., 2007). (*3*) Finally, *cheD* and other non-*che* genes, not found in the *E. coli* chemotaxis cluster, are present in the *E. coli*-like *Geobacter* Type 2 clusters. By analogy to the functions assigned in *B. subtilis* and *T. maritima*, the presence of CheD signifies that a specific mechanism is in play for deamidating MCPs and regulating CheY~P hydrolysis (Chao *et al.*, 2006, Rosario & Ordal, 1996). The presence of genes with unknown functions within chemotaxis operons has been reported in various bacteria, and appears to be commonplace in bacteria with more complex chemotaxis pathways (Butler & Camilli, 2005, Charon & Goldstein, 2002, Marchant *et al.*, 2002, Stover *et al.*, 2000).

Based on these observations, we suggest that the protein complexes encoded by Type 1 *che* clusters function like the *E. coli* chemotaxis pathway, albeit with the differences noted above, in which case *che* cluster 1 of *G. metallireducens* regulates signaling through a lone class 36H MCP (Gmet_1078) [34]. If cluster 2 of *G. uraniireducens* serves a similar role, then a lone class 36H MCP (Gura_2167) serves to detect the environmental stimuli in this situation as well. According to this reasoning, we do not expect *G. sulfurreducens* PCA (AE017180.1) to have a flagellum-controlling chemotaxis pathway that uses this signaling logic, because it lacks both a Type 1 *E. coli*-like *che* gene cluster and class 36H MCPs. However, the absence of class 36H MCPs does not rule out other modes of flagellum-controlling chemotactic signaling. For example, the chemotaxis pathway in *B. subtilis* uses class 44H MCPs, and the genome of *G. sulfurreducens* contains several *mcp* genes that belong to this class. Fewer investigations of Type 2 *E. coli*-like *che* clusters have been conducted, but in their study of a Type 2-like cluster in *R. sphaeroides*, Armitage and colleagues found that this cluster

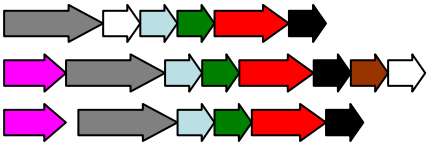
is essential for flagellar motility (Porter *et al.*, 2002). All the *Geobacter* species contain at least one Type 2 cluster; these too could potentially participate in flagellum-controlling chemotaxis. Further work is needed to verify the actual functions and relationships of Type 1 and 2 *E. coli*-like clusters, which will serve to strengthen the confidence of predictions based on percent gene identity, gene cluster organization and mechanistic similarities reflected in protein organization.

Dif-Like Clusters May Regulate Extracellular Matrix Formation and Chemotactic Motility

G. sulfurreducens and *G. uraniireducens* possess clusters comprised of similar genes and gene ordering to the *dif* cluster of *M. xanthus*. These clusters contain class 44H MCPs with two predicted transmembrane segments, but small periplasmic domains (~3 to 10 aa), genes for CheA, CheW, CheY, CheC and CheD, and genes of unidentified function (See Figure 9 for gene arrangements and percent identities of the individual gene products). The *dif* signaling system of *M. xanthus* has been studied most, where it is known to be involved in the regulation of exopolysaccharide formation, an essential component of the *Myxococcus* social motility apparatus (Black et al., 2006, Black & Yang, 2004). It has been noted that the phenomenon of social motility in *M. xanthus* resembles biofilm formation in other bacteria (O'Toole *et al.*, 2000). In addition, the *dif* cluster is involved in sensing certain lipids (Bonner et al., 2005). One difference between the *M. xanthus dif* cluster and the *Geobacter dif*-like clusters is the presence of *cheR* in the *Geobacter* cluster instead of *difB*. A plausible consequence of this observation is that the *Geobacter dif*-like pathways are CheR-dependent, whereas the *M. xanthus dif* system is CheR-independent. The involvement of *Geobacter dif*-like clusters in the synthesis of

extracellular matrix material, which is essential for biofilm formation, is currently under investigation.

A. Gene arrangement



B. Percent identities of full-length gene products in the *Geobacter dif*-like clusters to *M. xanthus dif*-homolog proteins. (The *E. coli* homologs are indicated in parentheses; c-DifA refers to the c-domain of the *M. xanthus* MCP.)

M. xanthus	**c-DifA (MCP)**	**DifB**	**DifC (CheW)**	**DifD (CheY)**	**DifE (CheA)**	**DifG (CheC)**
G. sul (cluster 6)	38	No homolog	18	54	43	33
G. ura (cluster 1)	38	No homolog	21	52	41	31
Mean identities with remaining *Geobacter* clusters[a]						
G. sul	27 ± 3 (7)		17 ± 2 (6)	28 ± 4 (5)	28 ± 4 (3)	No homolog
G. ura	23 ± 4 (5)		16 ± 4 (7)	24 ± 6 (8)	28 ± 4 (6)	No homolog

[a]mean ± standard deviation (number of homologs in average).

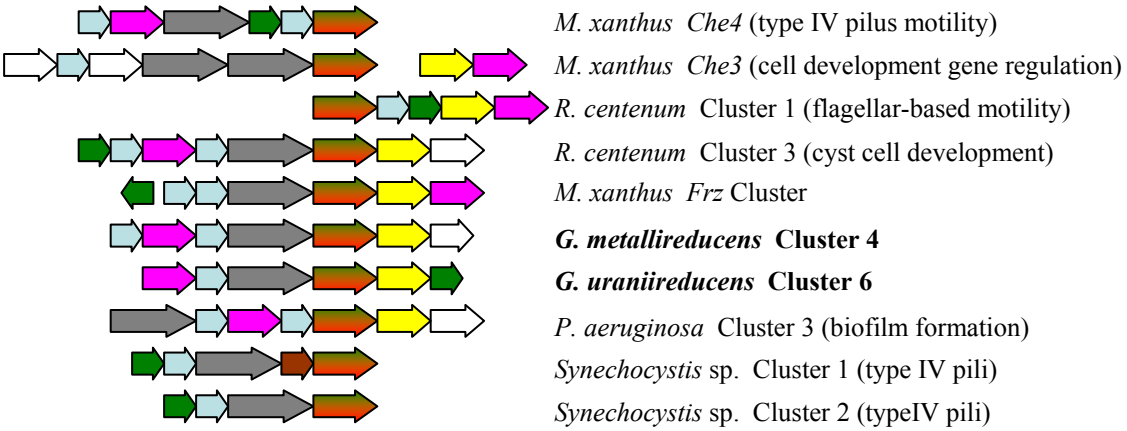
Figure 9. Gene arrangement of the *dif*-like clusters in *Geobacter* species (A) and the percent identity of the individual gene products in the cluster (B)

Che Clusters with CheA/Y Fusion Proteins

G. metallireducens and *G. uraniireducens* each have one *che* cluster with a gene that encodes a CheA-CheY fusion protein (CheA/Y). In *R. centenum*, *M. xanthus, P. aeruginosa*, and *Synechocystis* strain PCC6803, *cheA/Y*-containing *che* clusters carry out various functions, including the regulation of flagellar motility (Jiang & Bauer, 1997, Jiang *et al.*, 1997), type IV pilus-based motility and/or the biogenesis of type IV pili (Bhaya et al., 2001, Blackhart & Zusman, 1985, Vlamakis et al., 2004), cell development (Berleman & Bauer, 2005b, Kirby & Zusman, 2003), and biofilm formation (Hickman et

al., 2005). The *Geobacter che* clusters in Figure 3 that encode CheA/Y fusion proteins are most similar to the *M. xanthus* Frz cluster, cluster 3 of *P. aeruginosa*, and cluster 3 of *R. centenum* (by gene cluster content, gene order and the percent identity of CheA/Y homologs, see Figure 10), clusters that function in developmental cell aggregation (Li et al., 1997), biofilm formation (Hickman et al., 2005), and cyst cell development (Berleman & Bauer, 2005a), respectively – processes that involve cell-cell interaction. By these same criteria, the *Geobacter* clusters were least similar to *che* cluster 1 of *R. centenum* (chemotactic and phototactic responses (Jiang et al., 1997)), *M. xanthus* cluster 3 (regulation of fruiting body formation (Kirby & Zusman, 2003)) and *Synechocystis* cluster 2 (Bhaya et al., 2001). Overall, these findings suggest that the corresponding *Geobacter che* clusters may also regulate processes involving cell-cell interactions and/or social motility, but this is in need of experimental proof.

A. Gene arrangement



B. Percent identities between *Geobacter* CheAY fusions in other species.

Species and Cluster	**Protein Name**	**Identity (%)**	
		Gura_4171	**Gmet_2710**
M. xanthus Frz Cluster	FrzE	33	28
P. aeruginosa Cluster 3	WspE	30	29
R. centenum Cluster 3	CheA3	30	32
Synechocystis Cluster 1	TaxAY1	30	27
R. centenum Cluster 1	CheA1	29	26
M. xanthus Cluster 3	CheA3	27	29
Synechocystis Cluster 2	TaxAY2	25	24

Figure 10. Gene arrangements in *Geobacter* clusters containing *cheA/Y* fusions compared to other species whose functions have been identified (A), and the percent identities of *cheA/Y* fusions (B).

Che Clusters that are Unique to *Geobacter* Species and *δ-Proteobacteria*

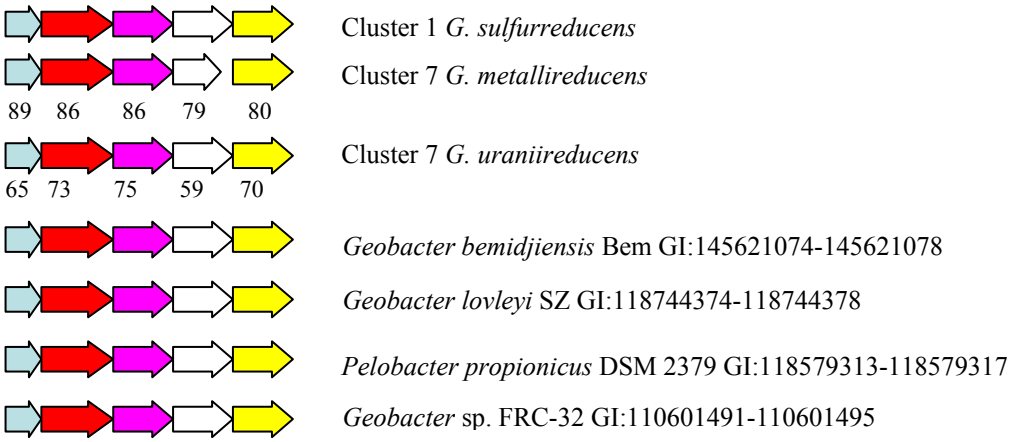
Two groups of *che* clusters are highly conserved among *Geobacter* species; we refer to these as α and β groups. The clusters belonging to these two groups contain the well-known homologs of chemotaxis genes (*cheA*, *cheW*, *cheB*, and *cheR*), but no *mcp* genes. Cluster 1 of *G. sulfurreducens* and cluster 7 of *G. metallireducens* and *G. uraniireducens* belong to the α group; the β group *che* clusters are 5, 2 and 5, respectively (Figure 3). An extensive search of both completed and draft bacterial genomes led to the finding that group α *che* gene clusters are present only in the genomes of the *Geobacteraceae*, including *Geobacter bemidjiensis* Bem, *Geobacter lovleyi* SZ, *Pelobacter propionicus* DSM 2379, and *Geobacter* sp. FRC-32.

Group α clusters have not been found in genomes outside the *Geobacteraceae* family (See Figure 11 for gene arrangements and percent identities). Each group α cluster contains a gene encoding a protein with an HD domain – which defines membership in an enzyme superfamily of metal-dependent phosphohydrolases, where the conserved His-Asp (HD) doublet has a role in catalysis (Aravind & Koonin, 1998). Within a variety of contexts, HD-domain-containing proteins have diverse biochemical functions, including nucleic acid metabolism and signal transduction. The predicted *Geobacter* homologs contain no other recognizable domains, *i.e.* they may function as standalone proteins. Standalone HD domain proteins in *E. coli* have low amino acid identity with each other and to the *Geobacter* homologs (~10%), and yet the *E. coli* proteins all act on nucleotide substrates (Zimmerman *et al.*, 2008). The predicted HD domain proteins located within the group α *che* clusters are probably regulated by, or participate in, chemotaxis-like signaling pathways of special significance to the cellular physiology of *Geobacter*. Our

preliminary data suggest that *Geobacter* species use this unique chemotaxis cluster to regulate chemotaxis (Chapter 4).

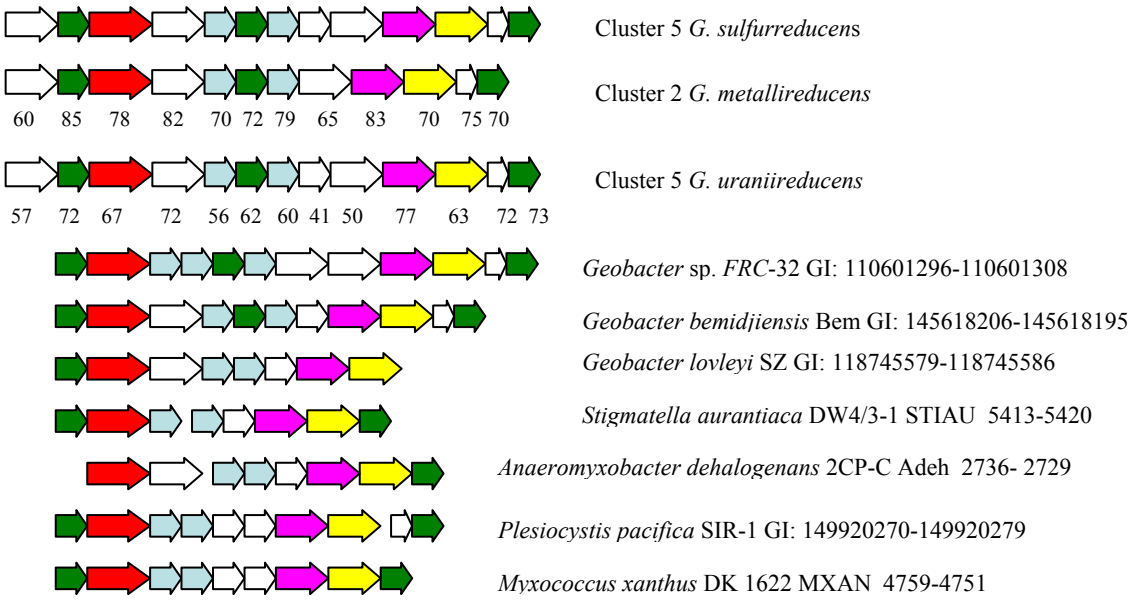
Group β clusters are conserved among δ-*Proteobacteria*, and have been identified in the genomes of *G. bemidjiensis*, G. *lovleyi*, *Geobacter* sp. FRC-32, *Stigmatella aurantiaca* DW4/3-1, *Anaeromyxobacter dehalogenans* 2CP-C, *Plesiocystis pacifica* SIR-1, and *Myxococcus xanthus* DK 1622. However, the functions of β group clusters are not known. Our data, reported in the following chapter, showed that the *che5* cluster of *G. sulfurreducens* regulates the expression of extracellular materials that are important for biofilm formation.

Group α: *che* cluster conserved among *Geobacteraceae*



Percent identities of the full-length *G. metallireducens* and *G. uraniireducens* gene products compared to *G. sulfurreducens* are shown below the *G. metallireducens* and *G. uraniireducens* clusters. Clusters found in other *Geobacteraceae* are shown with the corresponding NCBI Gene Identifier numbers or locus tags on the right.

Group β: *che* cluster conserved among *δ-proteobacteria*



Numbers below each gene are percent identities of the full-length *G. metallireducens* and *G. uraniireducens* gene products compared to *G. sulfurreducens*. The clusters found in other *δ-proteobacteria* are shown with the corresponding NCBI Gene Identifier numbers or locus tags on the right.

Figure 11. Gene arrangement of other *che* clusters

Chemotaxis Gene Expression Regulated by Alternative Sigma Factors 28 and 54

The mechanisms that regulate the expression of chemotaxis and flagellar genes are complex, and diverse, and should provide clues to the diversity and purpose of chemotaxis-like signaling systems. Therefore, we conducted a preliminary investigation into the regulation of *che* gene expression, in particular σ^{28}- and σ^{54}-regulated promoters upstream of *che* and flagellar gene clusters. In *E. coli* and *Salmonella*, the *che* and late flagellar genes, including *fliC* (the flagellar filament), are positively regulated by σ^{28} (Chevance & Hughes, 2008, Macnab, 1996, Soutourina & Bertin, 2003). In other bacteria, especially those with more than one *che* cluster, expression is also regulated by σ^{54} (McCarter, 2006). For instance, *R. sphaeroides* has a σ^{28}-regulated system that shows coupled expression of the chemotaxis proteins and flagella, and a system that regulates flagellar synthesis independently *via* σ^{54} (Martin *et al.*, 2006).

We searched upstream of the major *Geobacter che* operons and *fliC* loci for evidence of σ^{28}-regulated expression. As Figure 12 shows, σ^{28}-binding sites were identified upstream of *fliC* in *G. metallireducens*, *G. sulfurreducens* and *G. uraniireducens*, but only one major *G. sulfurreducens* group α cluster (cluster 1, Figure 3), had a recognizable σ^{28} binding site. Therefore, it seems that the specific mechanisms of regulation for most of the *Geobacter che* clusters will be different from *E. coli* (and *Salmonella*).

The *G. sulfurreducens* genome was searched for σ^{54} recognition sites to determine the number of *che* gene-related sites relative to all the sites that may exhibit σ^{54} regulation. Of the 110 sites identified genome-wide, nine were located in noncoding regions upstream of *che*, *mcp* or flagellar operons (Figure 13 lists positions and sequences

of the chemotaxis-related promoter sites) – one of these was the *Dif*-like cluster (cluster 6, Figure 3). Focused searches upstream of the major *che* clusters in the other two species identified possible σ^{54}-regulated promoter sites before cluster 3 in *G. uraniireducens,* and clusters 1 and 6 in *G. metallireducens* (Figure 12). No correlation was apparent between the identity of these clusters and their mode of regulation, *i.e. G. metallireducens* cluster 3 is classified as an *E. coli*-like cluster, and the other two do not belong to any identified class. Consequently, little specific insight can be gleaned from these early findings. Nonetheless, the results may presage a diversity of mechanisms for regulating expression. Indeed, we can expect that once the *che* gene-specific regulatory elements are known (which is significant in itself), it will be a challenge to determine how these systems map onto the global patterns of gene expression; this pattern should reflect how *Geobacter* adapts to the complex environment it inhabits.

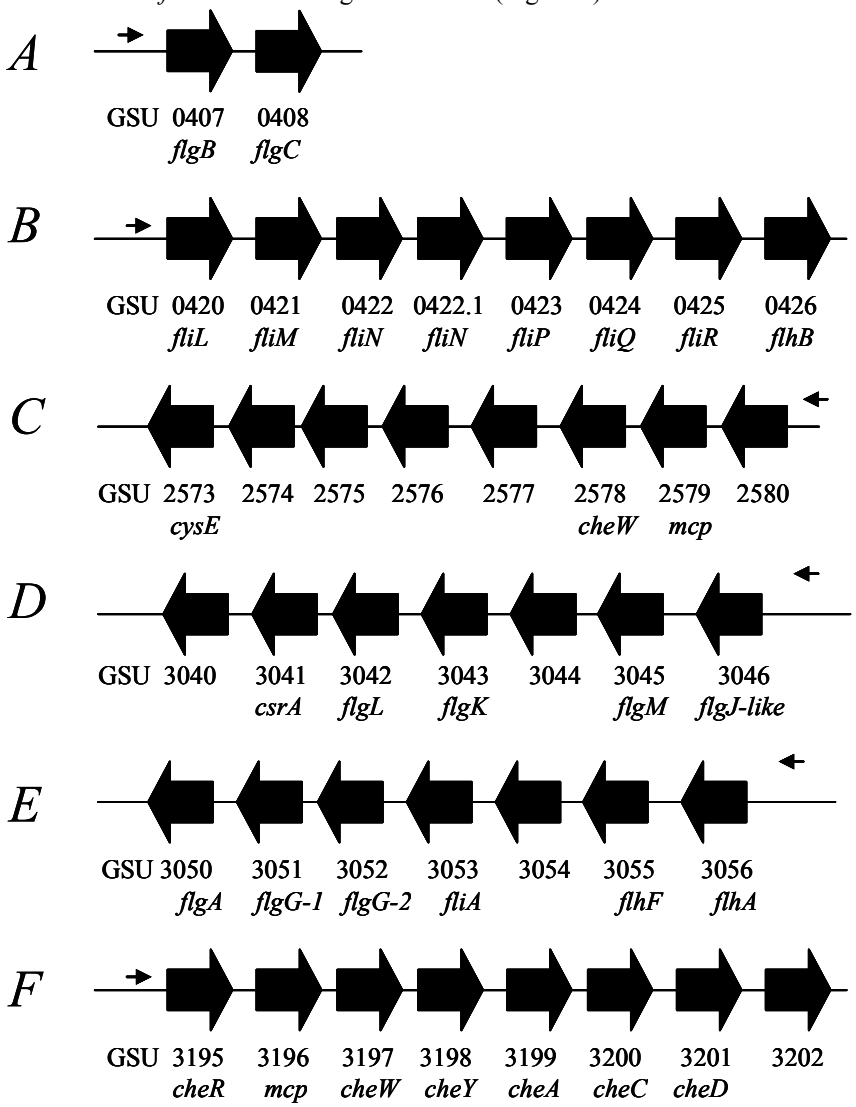
G. sul che cluster 1	AGGGGTACAAAAAAAATAC**TAAA**ATTATGGACTCTTGC**GCCGAT**G**A**GAATACTGCACCCG
gsu3038	GACGCAGCGATTTTTTTTC**TAAA**GCTTTTCCGGCTGCC**GCCGATA**CGTGAACTAAAAGGC
gura4096	ATTTTTTGACATTTTTTTC**TAAA**GTTTCTGCCGGCGCCT**CCGATAA**ATGAACTAACGGTC
gmet0442	GTAAATGTAATTTTTTTTC**TAAA**GTTTTTTTTGCCCCT**GCCGATAA**CGTTACCAAGAGCA
σ28 consensus	**TAAA**------15N------**GCCGATAA**

G. met che cluster 1	C**TGGCA**TTCCGGC**TGCA**
G. ura che cluster 3	G**TGGCA**TGCTCC**ATGCT**
G. met che cluster 3	CC**GGAAC**GGTTCT**TGCT**
G. sul che cluster 6	AC**GGAAC**ACTTC**TTGCT**
σ54 consensus	C**TGGCAC**-4N-**TTTGCA/T**

Figure 12. Putative σ^{28} and σ^{54} promoter elements. (A) Putative σ^{28} promoter regulation sites found upstream of *G. sulfurreducens che* cluster 1 and the *fliC* genes of *G. sulfurreducens* (*gsu3038*), *G. metallireducens* (*gmet_0442*), and *G. uraniireducens* (*gura_4096*) (Kutsukake *et al.*, 1990).The predicted transcription start sites are separated from the start codons by 125, 160, 127, and 152 nucleotide bases, respectively. (B) Putative σ^{54} promoter elements upstream of the *Geobacter* major *che* gene clusters (Barrios et al., 1999). For *G. metallireducens* clusters 1 and 3, *G. uraniireducens* cluster 3, and *G. sulfurreducens* cluster 6, the predicted transcription start sites are 50, 33, 24 and 16 nucleotide bases upstream of predicted operon ATG start codons, respectively.

Fig.	σ^{54}-binding sequence	Score	Position	Operon	Bases from 1st ATG
A	C**TGG**T**AC**GGCT**TTTGCT**	92	439442-439458	*Gsu0407-0408*	23
B	C**TGGCA**TTTCGG**TTGCA**	86	451031-451047	*Gsu0420-0426*	47
	A**TGGCAC**GGCCTG**TG**T**A**	81	773999-774015	*Gsu0725-0726*	26
C	T**TGGCA**TCCTGCC**TGCT**	81	2842557-2842541	*Gsu2573-2580*	48
	TC**GGCAC**GTAGG**TTGCA**	86	3311556-3311572	*Gsu3017-3028*	48
D	T**TGGCAC**ATAACA**TGCT**	86	3350142-3350126	*Gsu3040-3046*	48
E	C**TGGCAC**AACGG**TTGCA**	93	3359815-3359799	*Gsu3050-3056*	39
	AC**GGCAC**CGGCA**TTGC**C	80	3462928-3462944	*Gsu3156 (mcp)*	115
*F**	AC**GGAAC**ACTTC**TTGCT**	81	3503117-3503133	*Gsu3195-3202**	16
	CTGGCAC-4N-**TTTGCA/T** (*consensus*)				

Figure 13. Chromosomal locations of σ^{54} regulation elements in *G. sulfurreducens* upstream of *che* and flagellar gene operons
**Gsu3195-3202* is *G. sulfurreducens che* gene cluster 6 (Figure 3).



Conclusion

The comparative analysis of *che* gene clusters and regulatory sequences among *Geobacter* species and other bacteria has provided valuable insight into the functions of the various *Geobacter* chemotaxis-like pathways. The genomes of *Geobacter* species have multiple copies of chemotaxis genes – more than 60 genes per genome. Their arrangement in six to seven major clusters reflects both greater complexity and diversity in comparison to the single cluster on the *E. coli* chromosome. This diversity is also reflected in the presence of both σ^{54}- and σ^{28}-dependent regulatory sequences. The presence of multiple chemotaxis-like clusters and mechanisms of regulation both suggest that the pathways in *Geobacter* are not redundant, but instead each fills a specific cellular need.

Geobacter genomes have several chemotaxis-like clusters in addition to a *che* cluster that is similar in organization to the chemotaxis operons of *E. coli* and *S. enterica*. These clusters are similar to characterized clusters in other bacteria that regulate functions other than flagellar motility. From our analysis, it seems probable that *Geobacter* species use chemotaxis-like signaling pathways for a variety of functions, which probably include type IV pilus-based motility, regulation of motility apparatus expression (flagella, pili, and extracellular matrix), and biofilm formation Interestingly, the *Geobacter* species also have *che* clusters that – at the present time – appear to be unique, which may plausibly mean that these pathways regulate physiological functions that are unique to the genus *Geobacter*. Sensory inputs to the chemotaxis-like pathways are likely to be diverse, because the *Geobacter* genomes contain a large number of chemoreceptor (*mcp*) genes, which display a diversity of sensing domain architecture.

The presence of this large number of proteins – receptors and Che proteins – undoubtedly reflects a greater need for *Geobacter* species to respond to a variety of environmental conditions, which allows them to thrive in subsurface environments. The presence of *mcp* genes that belong to different MA domain classes in one genome – *i.e.* expression of MCPs in the same cell membrane with MA domains of different lengths, may imply segregation of receptors into class-specific clusters with their cognate Che signaling proteins. We postulate that this mechanism will generate pathway specificity and diminish unwanted cross-talk. Such a mechanism could be general for bacteria with multiple chemotaxis-like pathways.

CHAPTER 3

A CHE CLUSTER OF GEOBACTER SULFURREDUCENS REGULATES GENE EXPRESSION

Introduction

Geobacter species are Gram-negative δ-*proteobacteria* found predominantly in the Fe(III)-reduction zone of sedimentary environments. They can utilize organic compounds as electron donors and metal compounds or anodes of microbial fuel cells (MFCs) as electron acceptors (Lovley, 2008b). *Geobacter* species have been used for bioremediation of uranium-contaminated groundwater and have the potential to remediate organic wastes and generate electricity (Lovley, 2008b). *Geobacter sulfurreducens* has been extensively studied for better understanding of molecular processes that are involved in bioremediation and electricity generation, and to enhance the applications of *Geobacter* species.

Under laboratory conditions, *G. sulfurreducens* cells are planktonic and non-motile. When grown in MFCs for electricity generation, cells form a red-brown conductive biofilm surrounding the electrode. Biofilm formation and outer membrane *c*-type cytochromes OmcS and OmcZ have been shown to be important for *G. sulfurreducens* to produce optimal current (Lovley, 2008a, Lovley, 2008b). Strains that are defective in biofilm formation and/or bear deletions of *omcS* or *omcZ* exhibit diminished power production (Lovley, 2008b). To grow on insoluble electron acceptors such as Fe(III) oxide, an environment that the cells mostly encounter in nature, bacteria have to transfer electrons from inside the cells to the outside. OmcS has been shown to be essential for these processes in *G. sulfurreducens* (Mehta *et al.*, 2005). *G. sulfurreducens*

appears to control the expression of OmcS efficiently. OmcS is needed and expressed when cells grow on insoluble Fe(III) oxide, but not when they are grown with soluble Fe(III), and the level of *omcS* transcription increases as current production in MCF increases (Reguera *et al.*, 2005). How *G. sulfurreducens* regulates the expression of OmcS is unknown.

Chemotaxis is a trait shared by many bacteria that enables cells to move toward chemical attractants and away from repellents. The Che pathway of *E. coli*, which controls cell movement by regulating flagellar rotation, has been well-characterized. There are 11 key proteins of the *E. coli* Che pathway, including five methyl-accepting chemotaxis proteins (MCPs), an autophosphorylating histidine kinase (CheA), a scaffold protein (CheW), a methyltransferase (CheR), a methylesterase (CheB), a response regulator (CheY), and CheY phosphatase (CheZ). With the exception of three *mcps*, all remaining *che* genes are in one cluster. Chemotaxis homologues are found in most genomes of motile bacteria, and often in greater number, with multiple copies of each *che* gene arranged in multiple clusters (Antommattei & Weis, 2006, Tran *et al.*, 2008). A significant number of studies on species with multiple *che* clusters have demonstrated that *che* gene products of the same cluster work together, form a separate pathway, and carry out specific cellular functions. Often, one *che* cluster in a species with multiple *che* clusters regulates chemotaxis, and other *che* clusters regulate diverse functions such as biofilm formation, biosynthesis of the motility apparatus, or other gene expression (Tran et al., 2008).

The genome of *G. sulfurreducens* contains ~ 70 *che* and *mcp* homologs, arranged in 6 major clusters (Chapter 2). We investigated the functions of *che* genes in *G.*

sulfurreducens by making single deletion mutants of four *cheAs* (*gsu0296*, *gsu1290*, *gsu2222*, and *gsu3199*). Under laboratory conditions, deletion of *gsu2222* enhances cell aggregation. *Gsu2222* locates in the *che5* cluster, which consists of a complete set of genes coding for chemotaxis signaling proteins, including a *cheA*, a *cheB*, a *cheR*, 2 *cheWs*, and 3 *cheYs*, together with non-chemotaxis genes. Unlike most *che* clusters, in which at least one *mcp* is found, the *che5* cluster has no *mcp*. *Che*5 is classified as β-group, which is found exclusively in δ-*proteobacteria*, and the functions of clusters in this group are unknown (Tran et al., 2008). We investigated the function of *che5* by making deletions of individual genes in the cluster, and of the whole cluster. Our data indicate that *che5* of *G. sulfurreducens*, possibly together with one MCP class (40+24H), is involved in expression regulation of two important outer membrane *c*-type cytochromes, OmcS and OmcZ, and of genes possibly involved in biofilm formation. A model for Che5 pathway regulation of gene expression is proposed.

Methods

Strains, plasmids, and growth conditions

The plasmids and *G. sulfurreducens* strains used in this study are listed in Table 4. *E. coli* strain TOP10 [F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lacZ*ΔM15 Δ*lacX74 recA1 deoR araD139* Δ(*ara-leu*)*7697 galU galK rpsL* (Str) *endA1 nupG*] (Invitrogen Co., Carlsbad, CA) was used for DNA manipulations and PCR product subcloning. *G. sulfurreducens* strains were routinely cultured in acetate-fumarate medium at 30°C under strict anaerobic conditions as previously described (Lovley & Phillips, 1988a, Coppi *et al.*, 2001).

Designation	Genotype or description	Source or reference
Strains		
DL1MA	Wild type, originally called DL1	(Caccavo *et al.*, 1994a)
Δ*gsu2210*	DL1MA, *gsu2210::kan*	This study
Δ*cheY5a*	DL1MA, *gsu2212::kan*	This study
Δ*gsu2213*	DL1MA, *gsu2213::kan*	This study
Δ*cheB5*	DL1MA, *gsu2214::kan*	This study
Δ*cheR5kan*	DL1MA, *gsu2215::kan*	This study
Δ*cheR5spec*	DL1MA, *gsu2215::spec*	This study
Δ*gsu2216*	DL1MA, *gsu2216::kan*	This study
Δ*gsu2217*	DL1MA, *gsu2217::kan*	This study
Δ*cheW5a*	DL1MA, *gsu2218::kan*	This study
Δ*cheY5b*	DL1MA, *gsu2219::kan*	This study
Δ*cheW5b*	DL1MA, *gsu2220::kan*	This study
Δ*gsu2221*	DL1MA, *gsu2221::kan*	This study
Δ*cheA5*	DL1MA, *gsu2222::kan*	This study
Δ*cheY5c*	DL1MA, *gsu2223::kan*	This study
Δ*gsu2224*	DL1MA, *gsu2224::kan*	This study
Δ*gsu2225*	DL1MA, *gsu2225::kan*	This study
Δ*cheY5a*Δ*cheY5b*	DL1MA, *gsu2212::spec, gsu2219::kan*	This study
Δ*cheR5kan*/pRG5	Δ*cheR5*kan carrying pRG5	This study
Δ*cheR5kan* /pHT51	Δ*cheR5*kan carrying pHT51	This study
Δ*cheR5spec*/pCD341	Δ*cheR5*spec carrying pCD341	This study
Δ*cheR5spec*/pHT52	Δ*cheR5*spec carrying pHT52	This study
Plasmids		
pSJS985	Source of spectinomycin resistance cassette; $Spec^r$	(Sandler & Clark, 1994)
pRG5	Expression vector; $Spec^r$	(Kim *et al.*, 2005)
pCD341	Expression vector, and source of kanamycin resistance cassette; Kan^r	(Morales *et al.*, 1991)
pHT51	*gsu2215* in pRG5; $Spec^r$	This study
pHT52	*gsu2215* in pCD341; Kan^r	This study

Table 4. *G. sulfurreducens* strains and plasmids used in chapter 3

DNA manipulation and plasmids constructions

G. sulfurreducens genomic DNA was purified using the MasterPure Complete DNA purification kit (Epicentre Technologies, Madison, WI.). Mini Plasmid purification kits, PCR purification kits, and Qiaquick gel extraction kits (QIAGEN Inc. Valencia, CA) were used for plasmid DNA purification, PCR product purification, and gel extraction respectively. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs, Inc. (Beverly, MA). Primers were purchased from Sigma-Genosys (Sigma-Aldrich Co. St. Louis, MO) and Operon Biotechnologies, Inc. (Huntsville, AL). Taq DNA polymerase (QIAGEN Inc., Valencia, CA), and Phusion polymerase (Finnzymes) were used for PCR amplifications.

To construct plasmids pHT51 and pHT52, coding sequence of *gsu2215* were amplified with primers HT167F and HT167R (SI Table 1), creating EcoRI and HindIII restriction sites, using the following conditions: 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 58°C for 30 s, and 72°C for 2.5 min and a final extension at 72°C for 10 min. Double digested (EcoRI and HindIII) PCR product of *gsu2215* and pRG5 or pCD341 were ligated and transformed into *E.coli* TOP10. The plasmids were purified and sequenced in University of Massachusetts sequencing facility; those with correct sequences were used for down stream processes.

Construction of *G. sulfurreducens* mutant strains

Genes in the *che5* cluster were individually disrupted by an antibiotic resistance cassette using a single-step gene replacement method as previously described (Coppi *et al.*, 2001), with minor modifications. To replace a gene of interest with an antibiotic resistance cassette, a linear DNA fragment (~ 2.2 kb) containing the antibiotic resistance

marker flanked by ca. 0.5 kb of sequence upstream and downstream of the gene was generated by recombinant PCR (Coppi et al., 2001) or by two-step ligation. Briefly, the upstream, downstream, and antibiotic resistance cassette sequences were amplified with corresponding primers listed in Appendix B. Following PCR recombination or ligation of the upstream fragment with the antibiotic resistance cassette and then the downstream fragment, the final ~ 2.2 kb fragments were gel-extracted. They were then amplified with the distal primers (forward primer of upstream, and reverse primer of downstream fragment), gel-purified, concentrated, and transformed by electroporation into *G. sulfurreducens* as previously described (Coppi et al., 2001). Mutants were confirmed by PCR.

Enrichment of Extracellular matrix (ECM) material

An ECM enrichment method described previously (Chang & Dworkin, 1994, Curtis *et al.*, 2007) was modified for liquid culture of *G. sulfurreducens*. Briefly, a stationary phase culture of *G. sulfurreducens* was harvested at 10,000 x *g* for 10 min at 10^{o}C. The cell pellet was resuspended in ¼ volume of TNE buffer (10 mM Tris [pH 7.5], 100 mM NaCl, 5 mM EDTA) (Chang & Dworkin, 1994), and transferred to a 30 mL Wheaton homogenizer to declump the cells gently. Sodium dodecyl sulfate (SDS) from a 10% stock was added to the cell suspension to a final concentration of 0.5%, and stirred for 5 min at room temperature. The mixture was centrifuged at 12,000 x *g* for 10 min at 10^{o}C. The pellet was washed five times in the same volume of TNE buffer to remove SDS. The final pellet was resuspended in 1/20 volume of 10 mM Tris HCl buffer (pH 7.5) containing 5x Complete EDTA-free protease inhibitor cocktail (Roche), homogenized by drawing it up and down through a syringe several times with the

18G1/2, 23G1, and then 25G7/8 needles (Becton Dickinson & Co., Franklin Lakes, NJ), and stored at -20^{o}C until use.

Filament preparations

Filaments were purified by the method for surface pilin preparation described by Alm and Mattick (Alm & Mattick, 1995), with some modifications. Briefly, a stationary phase culture of *G. sulfurreducens* was harvested at 5,000 x *g* for 5 min at 4^{o}C. The pellet was resuspended in 1/10 volume of 10 mM Tris HCl buffer (pH 7.5). The cell suspension was vortexed at maximum speed with a table top vortexer for 14 min to shear off the filaments, then centrifuged at 5,000 x *g* for 8 min at 4^{o}C to separate the cells from the filaments. The supernatant was passed through a 0.2 μm sterile filter to remove remaining cells, and then concentrated by filtering through a YM-30 Centriplus centrifugal filter device (Milipore Co., Bedford, MA), following manufacturer's instructions, until the retentate volume was about 0.5-1.0 mL. The retentate was centrifuged at 13,000 RPM for 20 min at 4^{o}C. The red pellet was collected and resuspended into 1/100 volume of 10 mM Tris HCl buffer (pH 7.5) containing 5x Complete EDTA-free protease inhibitor cocktail (Roche), and stored at -20^{o}C until use.

Electron microscopy

Approximately 5 uL of prepared filaments were adsorbed to freshly discharged carbon-coated copper grids (No. 200) for 2 min, then briefly blotted onto Whatman filter paper. The grids were then quickly placed on a 20 μl drop of 2% (w/v) uranyl acetate (pH 3.9) for 45 s and then blotted with Whatman filter paper to dry. The grids were examined, and images were taken using a JEOL 100S transmission electron microscope operated at 80 V.

Samples for scanning electron microscopy (SEM) were prepared essentially as previously described (Bond & Lovley, 2003), with the following exceptions. Biofilm cells were obtained by growing *G. sulfurreducens* anaerobically in a Petri dish containing glass cover slips. After three days, the cover slips were carefully removed from the Petri dish, and subjected to further treatment as described by Bond & Lovley (2003).

Gel electrophoresis, heme staining and Western blot analysis

Samples were electrophoresed on 12% Tris-Tricine gel (Ausubel, 1999) and visualized via Coomassie Blue staining. For Western analysis, samples after electrophoresis were transferred to a Polyvinylidene Fluoride membrane (Bio-Rad) using a semi-dry electroblotting system (Bio-Rad) per manufacturer's instructions. The presence of PilA on the membrane was detected using a one-step western kit from GenScript (GenScript Corp., Piscataway, NJ) following manufacturer's instructions with a primary antibody specific for PilA of *G. sulfurreducen*s. For detection of heme-binding proteins, samples were electrophoresed on 15% Next gel (Amresco), and stained with *N*,*N*,*N*',*N*'-tetramethylbenzidine as previously described (Francis & Becker, 1984, Thomas *et al.*, 1976). SeeBlue Plus prestained protein standards (Invitrogen) were used as markers for all electrophoresis experiments.

Protein quantification and identification

Protein concentrations were determined by the bicinchoninic acid (BCA) method with bovine serum albumin as a standard (Smith *et al.*, 1985).

For identification, protein bands of interest were excised from Coomassie Brilliant Blue-stained gels, then digested 'in-gel' as described elsewhere (Lahm & Langen, 2000). Digested samples were further purified via micro Zip Tipping. Briefly, 10

ul volumes were acidified with 1-2 ul of 5% trifluoroacetic acid (TFA). Samples were loaded on an uC18 Zip Tip (Millipore, Corp) after pre-equilibration in 0.1% TFA. After washing twice with 10 ul aliquots of 0.1% TFA, samples were deposited directly onto the MALDI sample target, followed by addition of 0.5 ul of Matrix solution (5 mg/ml of alpha-cyano-4-hydroxycinnamic acid in a 1:1 mixture of acetonitrile and 0.1% TFA). Samples were allowed to air dry prior to insertion into the mass spectrometer. Analyses were performed on a Shimadzu Biotech Axima TOF^2 (Shimadzu Instruments) matrix-assisted-laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer. Peptides were analyzed in positive ion Reflectron mode. The instrument was externally calibrated using a local spot to the sample of interest with Angiotensin II (1046.54), Angiotensin I (1296.69 Da), Neurotensin (1672.92), ACTH (1-17) 2093.09 and ACTH (18-39) 2465.20 Da. Collisionally Induced Dissociation (CID) analyses were performed on the same instrument using a dual timed ion gate for high-resolution precursor selection with a laser power about 20% higher than for MS acquisition and He as the collision gas. CID fragments were separated in a Curved Field Reflectron (CFR), which allowed for a seamless full mass range acquisition of the MS/MS spectrum. All spectra were processed with Mascot Distiller (Matrix Sciences, Ltd.) prior to database searching. Database searches were performed in house with Mascot (Matrix Sciences, Ltd.). For MS searches the Peptide Mass Fingerprint program was used with a peptide mass tolerance of 150 ppm. For MS/MS searching (CID spectra) the MS/MS Ion Search program was used with a Precursor tolerance of 150 ppm and a fragment tolerance of 1 Da. Sample digestion and analysis were carried out at the University of Massachusetts – Medical School, Laboratory for Mass Spectrometry, Worcester.

Aggregation assay

Cells grown to stationary phase in 20 mL pressure tubes were harvested. Planktonic cells (2 mL), and 1 mL containing resuspended cells that were stuck to the side and bottom of the tubes were centrifuged at 8,000 x *g* for 5 min, and resuspended in 200 uL 10 mM Tris HCl buffer (pH 7.5). Total proteins were measured for the planktonic and sticky cell fractions using the BCA assay described earlier (Smith et al., 1985). The reported percentage of cells aggregated was the ratio between total protein in the sticky fraction divided by total protein in the tube (planktonic and sticky fractions), multiplied by 100.

Relative quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)

Total RNA was purified from mid-log cultures using RNeasy Midi kits (QIAGEN Inc., Valencia, CA). Purified RNA was then treated with the DNA-*free*™ kit (Ambion) following manufacturer's instructions to remove all contaminated DNA. RNA samples were used as template for PCR to check for DNA contamination. The quality and quantity of the RNA (DNA-free) were checked by visualizing on agarose gels and a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE) respectively, and stored at -80^{o}C until use. cDNA was synthesized using the Enhanced Avian HS RT PCR Kit (Sigma), primed with random nonamers (Holmes *et al.*, 2004). Primers for relative quantitative RT-PCR of genes of interest were designed for PCR fragments of ~ 100-130 bp in length, and checked for quality by visualizing PCR products on agarose gels before use in qRT-PCR. The selected primers are listed in Appendix B. For relative qRT-PCR, all cDNA reactions were made from the same concentration of RNA (0.1 μg/μL). Two microliters of cDNA were added per well of a

96-well optical plate containing 12.5 uL Power SYBR Green Master Mix (Applied Biosystems), 1.25 uL of 200 nM forward primer, 1.25 uL of 200 nM reverse primer, and 8 uL of RNase-free water. The reaction was run on a GeneAmp 7500 sequence detection system with GeneAmp 7500 SDS software (Applied Biosystems, Foster City, CA) using the following program: 1 cycle of 50°C for 2 min, then 1 cycle of 95°C for 10 min, then 40 cycles of 95°C for 15 s, and one cycle of 60°C for 1 min. Results were analyzed using GeneAmp 7500 SDS software, with normalization to the amount of cDNA in the wild type.

Microarray analysis

RNA samples (DNA-free) prepared from mid-log cultures of two biological replicates, divided into three technical replicates each, were concentrated using the sodium acetate/ethanol precipitation method, and resuspended in TE buffer to a final concentration of ~ 1 μg/μL, and then were subjected to microarray analysis using the method described in a previous study (Postier *et al.*, 2008). Data were analyzed using GenePix and Acuity 4.0 software as described earlier (Nevin *et al.*, 2009). Genes with significant differences (p value of ≤ 0.005) and log ratios greater than or equal to two were considered differentially regulated.

Results

Components and organization of the *che5* cluster of *G. sulfurredecens*

Che5 of *G. sulfurreducens* contains eight genes encoding chemotaxis proteins, including one CheA (*cheA5*), one CheR (*cheR5*), one CheB (*cheB5*), two CheWs (*cheW5a* and *cheW5b*), and three CheYs (*cheY5a*, *cheY5b*, and *cheY5c*). In addition, there are non-*che* genes found in the cluster (Figure 14a).

Many *che* operons contain at least one chemoreceptor. In *che5* of *G. sulfurreducens*, however, no *mcp* gene is found. Bioinformatic methods for operon prediction suggested that *che5* genes belong to an operon of 15 genes, from *gsu2224* to *gsu2210* (Tran et al., 2008). To determine if all of these genes are transcribed in a single transcript and whether other up/downstream genes are co-transcribed with them, we used RT-PCR. Primers were designed for PCR of the intergenic region between two constitutive genes (Appendix B). cDNA prepared from DNA-free RNA of *G. sulfurreducens*, as described in the Methods, was used as the template for PCR. As shown in Figure 14b, PCR products were detected in all intergenic regions from *gsu2225* to *gsu2210*, and not detected in the intergenic regions between *gsu2210-gsu2209* and *gsu2226-gsu2225*, suggesting that genes from *gsu2225* to *gsu2210* are co-transcribed.

There are 8 non-*che* genes present in the cluster, including *gsu2210*, *gsu2213*, *gsu2216*, *gsu2217, gsu2221, gsu2224* and *gsu2225*. GSU2210 is a *c*-type cytochrome with 27 putative heme-binding sites. GSU2213 contains a GAF (cyclic GMP, adenylyl cyclase, FhlA) domain (Hurley, 2003). *gsu2216* codes for a PBS lyase HEAT-like repeat protein (http://pfam.sanger.ac.uk/family?id=HEAT_PBS). GSU2217 contains a receiver domain. GSU2221 belongs to an AAA ATPase superfamily proteins. GSU2224 consists of a response receiver domain, together with a PATAN and a FRGAF domain (Makarova *et al.*, 2006). GSU2225 belongs to a family of GTPase EngA proteins. The domains of non-*che* gene products with putative functions are listed in Figure 15.

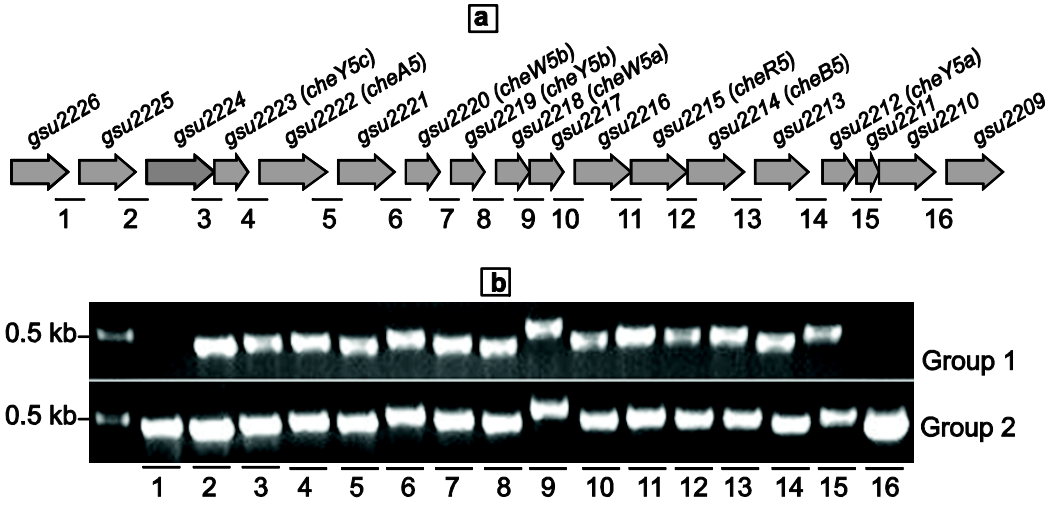


Figure 14. (a) Gene arrangement of the *che5* cluster. **(b)** RT-PCR to check co-transcription of *che5* cluster. Primers were designed to amplify ~ 400-500 bp covering intergenic regions between 2 constitutive genes, labeled from 1 to 16 for *gsu2226-gsu2209*. The total RNA (DNA free) of *G. sulfurreducens* was prepared to make cDNA with reverse transcriptase for use as template for PCR. If a PCR product is detected, two genes are co-transcribed. Group 1: PCR products using cDNA as template; group 2: PCR products using chromosomal DNA as template (positive control). The negative control (*not shown*), used RNA without RT as template, and no PCR product was detected, indicating that the RNA samples were free of contaminating genomic DNA. PCR products of fragments from 2-15 are present, but not for fragments 1 and 16, suggesting that *gsu2225-gsu2210* are co-transcribed.

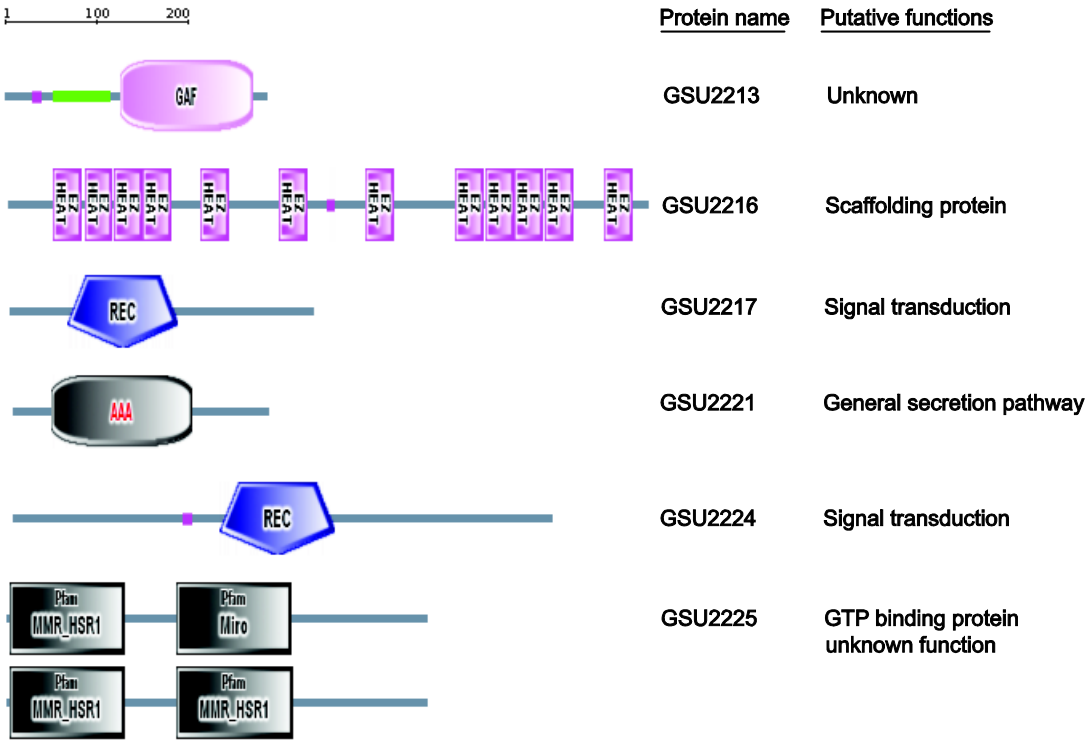


Figure 15. Domain prediction of non-Che proteins in the *che5* operon

Deletion of *cheR5*, *cheW5b*, and *cheA5* leads to enhanced cell aggregation and production of more filamentous materials on the cell outer surface compared to the WT

Under laboratory conditions (cells are grown in acetate-fumarate medium at 30^{o}C) the WT grows planktonically; however, Δ*cheA5*, Δ*cheR5* and Δ*cheW5b* cells tend to adhere to the bottom and sides of the culture tubes. Figure 16 shows the relative amount of cells that stuck to the culture tubes. There are 6, 9, and 11 times more cells aggregated and stuck to the culture tubes in *cheA5*, Δ*cheR5* and Δ*cheW5b* cultures, respectively, compared to the WT.

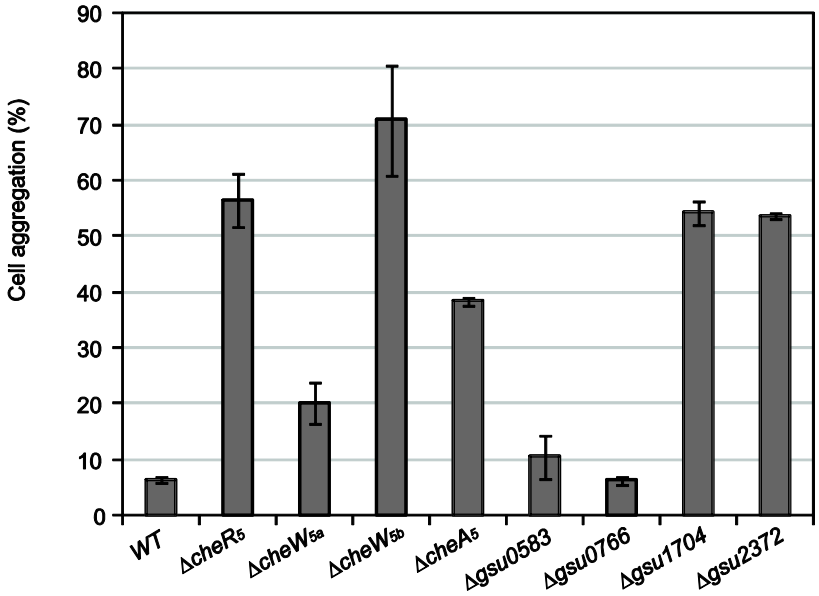


Figure 16. Aggregation assay of WT and *che* mutants (The bars are one standard deviation from the mean for three replicates.)

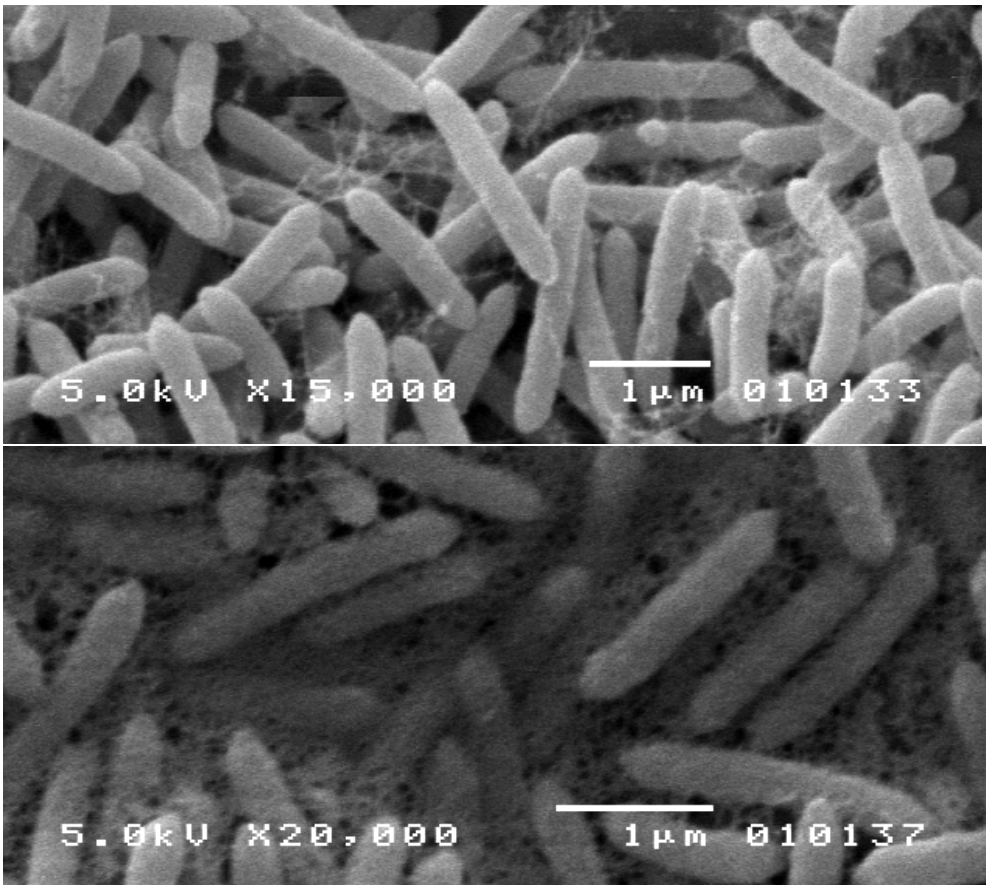


Figure 17. SEM image of the biofilm of the WT (top) and the Δ*cheR5* strain (bottom).

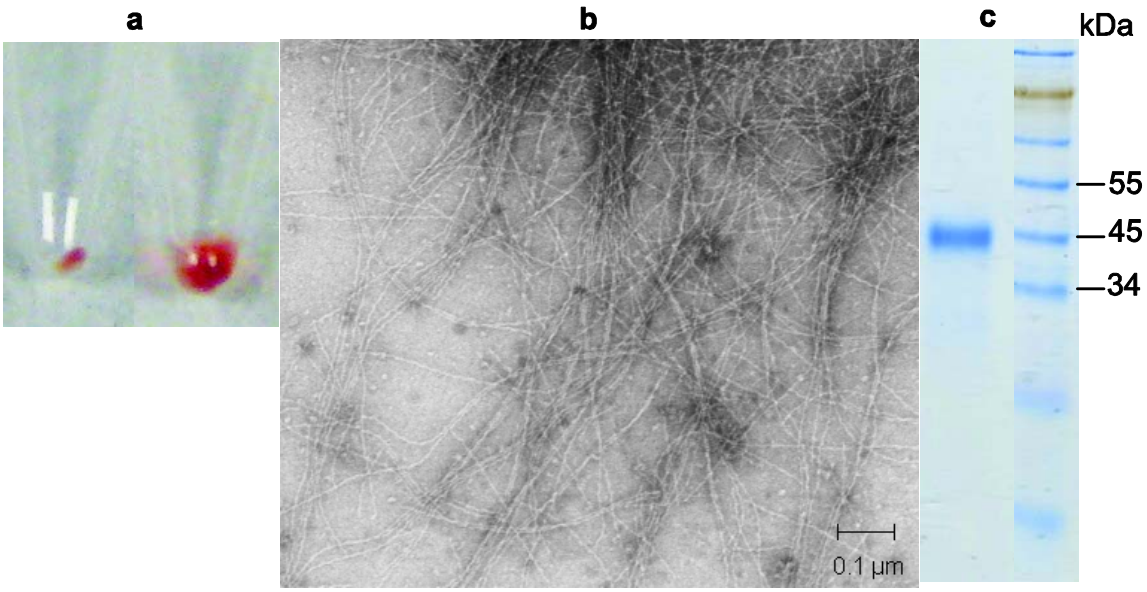


Figure 18. Filaments harvested from the WT and the Δ*cheR5* mutant. **(a)** Enriched filamentous materials from Δ*cheR5* are more abundant than in WT. **(b)** TEM images show that filamentous materials are over-produced in the Δ*cheR5* mutant (middle). **(c)** Electrophoresis of filamentous materials shows a predominant band at ~ 47 kDa.

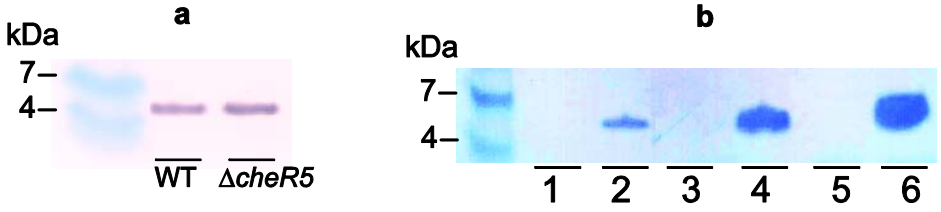


Figure 19. (**a**) Western blot of whole cells of the WT and Δ*cheR5* strains. Although more filaments are observed on a per-cell basis in the Δ*cheR5* strain compared to the WT, there is no distinguishable difference in the amount of PilA protein in Δ*cheR5* cells compared to the WT. **(b)** Western blot of filamentous materials after overnight proteinase K treatment (lanes 1, 3, and 5) and controls of filamentous materials without addition of proteinase K (lanes 2, 4, and 6). The amount of sample loaded was increased: lanes 3-4 and lanes 5-6 have 3.3 and 5 times more protein, respectively, than lanes 1-2.

We examined whether the cell surface of the WT is different compared to the Δ*cheA5*, Δ*cheR5* and Δ*cheW5b* mutants using transmission electron microscopy (TEM) for single cells, and scanning electron microscopy (SEM) for biofilm cells. By SEM, the biofilms of Δ*cheA5*, Δ*cheR5* and Δ*cheW5b* strains are indistinguishable; however, they are different from the WT. In all three mutants, greater filamentous materials in the extracellular matrix are seen compared to the WT (Figure 17). We also observed single cells (under TEM) of Δ*cheA5*, Δ*cheR5* and Δ*cheW5b* strains producing more filaments than the WT (data are not shown).

Filaments were purified by adapting a method for surface pilus preparation (see Methods), and characterized. As seen in Figure 18a, filmentous materials are more abundant in Δ*cheR5* cultures than in the WT, which was confirmed by observations of the product under TEM (Figure 18b). Similar results were achieved with Δ*cheA5* and Δ*cheW5b* strains (data not shown). The filamentous products were characterized by running them on a 10% Tris-Tricine gel, and staining with Coomassie Blue. As shown in Figure 18c there is a predominant band at the MW of ~ 47 kDa. To identify the protein, the band was excised and analyzed by mass spectrometry. A list of peptides detected is shown in Table 4. All of them match the sequence of GSU2504 (OmcS). The original hypothesis that OmcS makes up the filament, however, was ruled out, because after extended proteinase K digestion, no OmcS was detected by heme-staining of gels, but the filamentous structures were still observed (data not shown). Another hypothesis was that PilA makes up the filamentous structure, as stated in the previous report (Reguera et al., 2005). To test this hypothesis, filaments were digested overnight with proteinase K, checked by TEM, and run on a 10% Tris-Tricine gel for western analysis. There is no

distinguishable difference between the filaments seen by TEM of the overnight proteinase K-treated sample compared to the untreated sample (data not shown). Figure 19a show that PilA is not detectable in the proteinase K-treated filament, suggesting that the filamentous structure on the outer surface of *G. sulfurreducens* is composed of materials other than PilA. Other evidence indicats that the filamentous structures are not made up of PilA are: (a) that although more filaments are observed on a per-cell basis in the *cheR5* deletion mutant cells than in the WT, there is no distinguishable difference in the amount of PilA protein based on a western blot (Figure 19a); and (b) that there is no significant difference in transcription of *pilA* based on microarray analysis comparing the *cheR5* deletion mutant and WT (Appendix C). Filamentous materials are frequently observed at an elevated level in biofilm cells compared to planktonic cells of *G. sulfurreducens*, suggest their role in biofilm formation. Besides OmcS, other components of the filamentous materials are not yet determined, and are under investigation.

(K)FAPYQR(A)
(R)TEATTQTR(V)
(R)ILGGTGYQPK(S)
(R)AHASGFDSMTR(F)
(R)TADKFAPYQR(A)
(R)SVNEMTAAYYGR(T)
(R)SVNEMTAAYYGR(T)+ Oxidation (M)
(K)FGATIAGLYNSYK(K)
(R)FVDGSIATTGLPIK(N)
(K)FGATIAGLYNSYKK(S)
(R)RFVDGSIATTGLPIK(N)
(K)NSGSYQNSNDPTAWGAVGAYR(I)
(K)SLSGSYAFANQVPAAVAPSTYNR(T)

Table 5. Peptides from trypsin digestion of the major band in a Tris-Tricine gel of filament materials. The letters in parentheses indicate cleavage sites for the enzyme trypsin, which cleaves after K and R residues.

Extracellular materials isolated from single gene deletion mutants of *che5*

Extracellular materials (ECM) were prepared using an adaptation of the method for *M. xanthus* (see Methods). As seen in Figure 20b, there are two strongest bands in the heme-stained gel, which correspond to the molecular masses of OmcS and OmcZ (Izallalen *et al.* ASM 2008 Poster, Boston). Together with the Coomassie-stained gel (Figure 20a), this indicates that OmcS and OmcZ are two major cytochrome proteins on the cell surface of *G. sulfurreducens*. Figure 20 shows that deletion strains of Δ*cheA5*, Δ*cheR5*, Δ*cheW5a*, and Δ*cheW5b* overproduced OmcS while having a negative effect on OmcZ production compared to the WT. In contrast, Δ*cheB5*, Δ*cheY5c* and Δ*gsu2216* strains produced less OmcS and more OmcZ compared to the WT. Deletions of other non-*che* genes do not show significant effects on OmcS and OmcZ production (Figure 20).

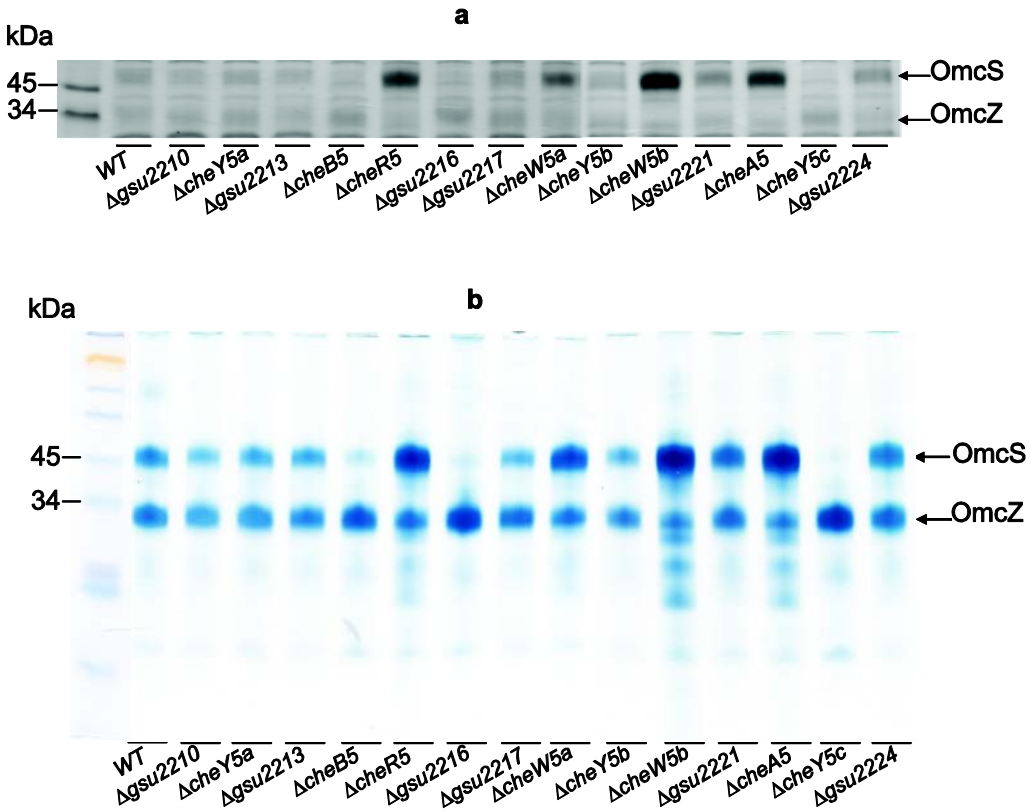


Figure 20. Coomassie **(a)**, and heme-stained **(b)** gels of ECM from single gene mutants in the *che5* operon: Δ*cheR5*, Δ*cheW5a,* Δ*cheW5b* and Δ*cheA5* mutants produced more OmcS and less OmcZ. In contrast, Δ*cheB5*, Δ*cheY5c*, and Δ*gsu2216* mutants produced less OmcS and more OmcZ than the WT.

Expression in *trans* of CheR5 in the Δ*cheR5* strain

To examine whether the phenotypes of the *che5* deletion mutants are in fact due to the absence of the gene and not due to polar effects on the downstream genes or unexpected secondary mutations, we conducted a complementation assay of one representative Δ*cheR5* strain. An expression vector containing the coding sequence of CheR5, pHT52, was introduced into Δ*cheR5*spec by electroporation, and single colonies were purified and grown in NBAF. In midlog phase, 0.05 mM-1.0 mM IPTG was added to the culture and growth was continued for 24 hours before harvesting cells for ECM preparation. Figure 21a shows the heme-stained gel of ECM from WT, Δ*cheR5*spec with pCD341 as control, Δ*cheR5*spec with pHT52 (uninduced) and Δ*cheR5*spec expressing *cheR5*. The results demonstrated that the presence of CheR5 in the strain lacking *cheR5* enabled the cells to recover OmcZ and reduce OmcS in the ECM to the WT level.

In another experiment, *cheR5* was introduced into the mutant Δ*cheR5*kan by a constitutive expression vector, pHT51. In *G. sulfurreducens*, this expression vector has been known to confer a higher copy number of inserted genes than the chromosome, and is used regularly for overexpression of genes of interest. For example, transcript of gsu2751 (dcuB), which was expressed in pRG5, is three times more abundant than the chromosome DL1 (0082-GSUL microarray data set- unpulished data). As shown in Fig21b, Δ*cheR5*kan/pHT51 produces more OmcZ and less OmcS than the WT, the same phenotype as that observed in the Δ*cheB5* strain.

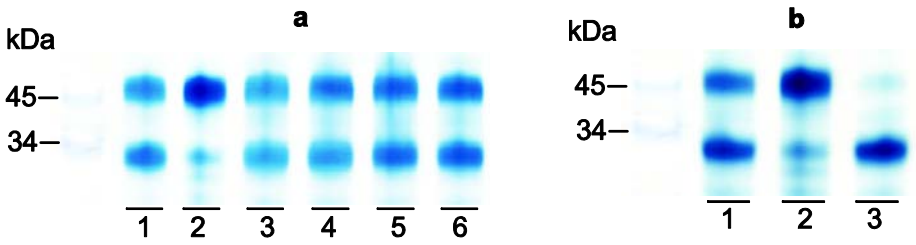


Figure 21. ***In trans*** **expression of CheR in Δ*cheR5* strains:** (**a**) Expression of CheR from pCD341 complements the Δ*cheR5* phenotype. Heme-stained gel of ECM from WT (lane 1) and Δ*cheR5*spec/pHT51 with 0 mM, 0.05 mM, 0.1 mM, 0.5 mM and 1.0 mM of IPTG (lanes 2 to 6, respectively). Strain Δ*cheR5*spec/pCD341 was used as control and does not show any difference from strain Δ*cheR5* (data not shown). (**b**) Expression of CheR5 from pRG5 (over-producing CheR5) shows a Δ*cheB5* phenotype. Heme-stained gel of ECM from WT (lane 1), Δ*cheR5*kan/pRG5 (lane 2), and Δ*cheR5*kan/pHT52 (lane 3).

Analysis of the transcription level of *omcS* and *omcZ* in the Δ*cheA5*, Δ*cheR5*, Δ*cheW5a*, Δ*cheW5b*, Δ*cheY5a*, Δ*cheY5b*, Δ*cheY5c*, Δ*cheB5,* Δ*gsu2216*, and Δ*gsu2224-2210* strains

To determine whether the changes in protein level of OmcS and OmcZ in *che5* mutants are due to regulation at the transcriptional level, we quantified the amounts of *omcS* and *omcZ* transcripts in mutants and WT using relative qRT-PCR. Total RNA was extracted from midlog cultures of WT, Δ*cheA5*, Δ*cheR5*, Δ*cheB5,* Δ*cheW5a*, Δ*cheW5b*, Δ*cheY5a*, Δ*cheY5b*, Δ*cheY5c*, and Δ*gsu2216*, and the whole-cluster deletion (*gsu2224-gsu2210*) strain Δ*che5*, grown in NBAF medium. cDNA was then synthesized from the same amount of RNA as described in Methods, using three sets of primers: HT224F and HT224R for *omcS*, HT216F and HT216R for *omcZ*, and HT219F and HT219R for *proC* (a housekeeping gene as a control), and quantified with normalization to the WT (Appendix B). As shown in Figure 22, the transcriptional levels of *omcS* in Δ*cheR5*, Δ*cheW5b,* Δ*cheA5*, Δ*cheW5a* strains are about 26, 18, 4, and 2-fold greater than the WT, respectively. Transcriptional levels of *omcZ* in Δ*cheR5*, Δ*cheW5b*, and Δ*cheA5* strains are 4, 3, and 2-fold lower than the WT, respectively (Figure 23). In contrast, strains of

Δ*cheB5*, Δ*gsu2216*, and Δ*cheY5c* showed decreases in *omcS* transcript levels (3, 67, and 91-fold, respectively), and increases in *omcZ* transcript levels of about 2-fold compared to the WT (Figures 22 & 23). In the *E. coli* chemotaxis pathway, deletion mutations of *cheA* and *cheR* lead to cells biased toward smooth swimming, and *cheB* deletion leads to an opposite phenotype (i.e., cells are tumble-biased). In this respect, the Che5 pathway appears to be similar to the *E. coli* chemotaxis pathway. A difference is that in *E. coli*, a deletion mutant of *cheY* shows a similar phenotype to deletion of *cheA* and *cheR*, but none of the *cheY* deletions in *che5* of *G. sulfurreducens* show the phenotype observed in *cheA5* or *cheR5* deletion mutant cells.

Global transcription analysis of the Δ*cheR5* strain

To examine whether the *che5* genes are involved in the regulation of other genes than *omcS* and *omcZ*, microarray analysis of the Δ*cheR5* strain compared to the WT was performed. Using the p value cut off of 0.005, there are 176 genes with significant changes (≥ 2 fold changes) in transcription, listed in Appendix C. Among these, 49 gene products are predicted to reside in the ECM according to the PSORTb program (Gardy *et al.*, 2005). OmcS and OmcZ are in this list (Table 6).

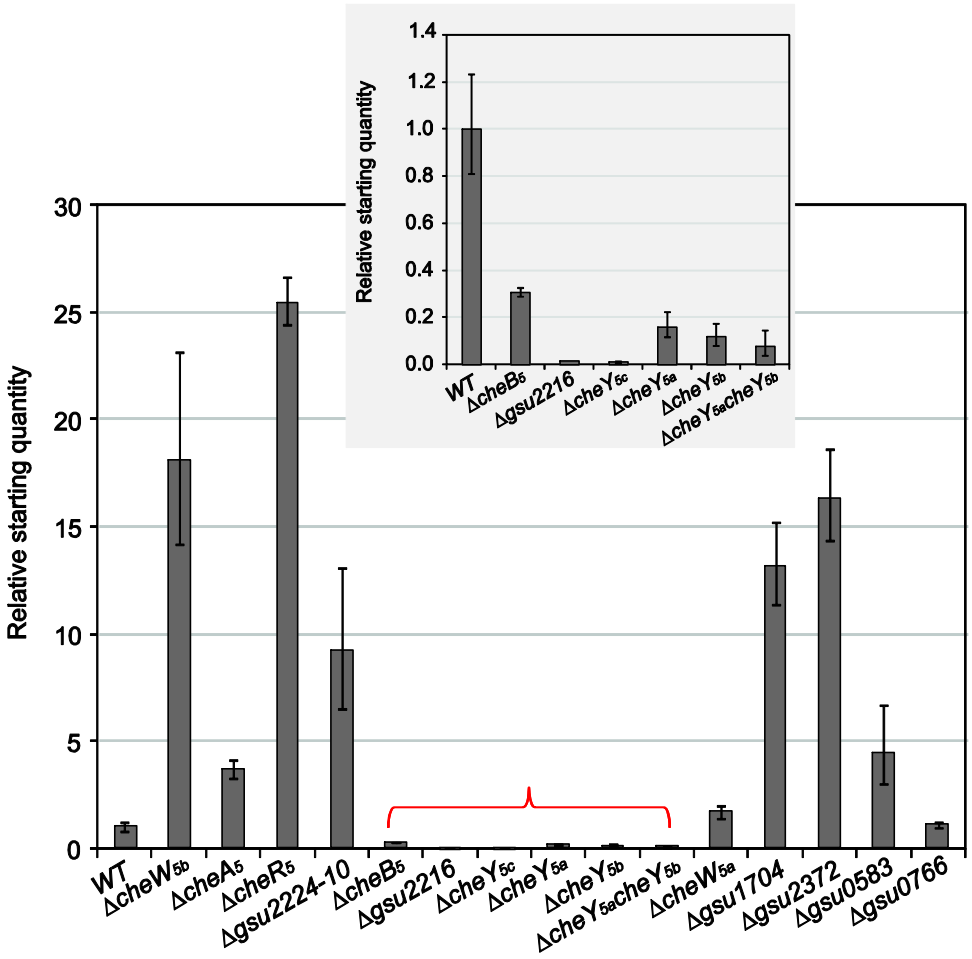


Figure 22. Transcription analysis by qRT-PCR for *omcS* in Δ*cheR5*, Δ*cheB5*, Δ*cheW5a*, Δ*cheW5b*, Δ*cheA5,* Δ*cheY5a*, Δ*cheY5b*, Δ*cheY5c*, Δ*gsu2216* and Δ*che5* strains. At least three technical replicates were used for each gene with each strain for statistical analysis. Error bars are based on 99.5% confidence.

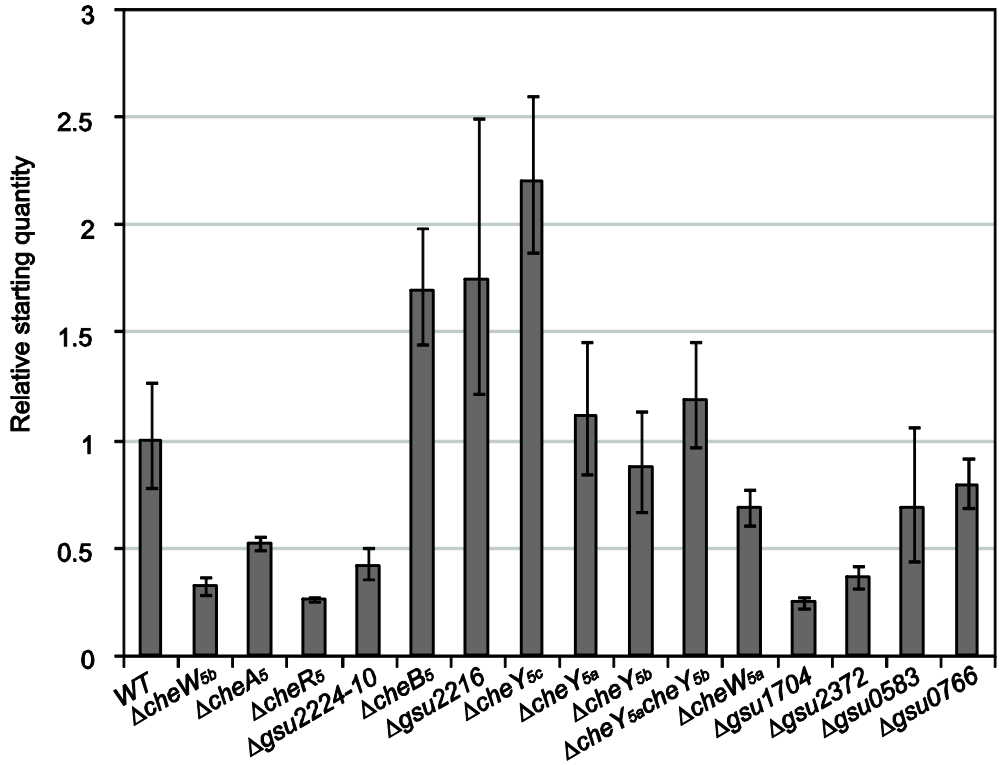


Figure 23. Transcription analysis by qRT-PCR for *omcZ* in Δ*cheR5*, Δ*cheB5*, Δ*cheW5a*, Δ*cheW5b*, Δ*cheA5,* Δ*cheY5a*, Δ*cheY5b*, Δ*cheY5c*, Δ*gsu2216* and Δ*che5* strains. At least three technical replicates were used for each gene with each strain for statistical analysis. Error bars are based on 99.5% confidence.

Probe Set ID	Fold change	Gene name	Annotation
GSU2503	11.3 up	*omcT*	cytochrome *c*, 6 heme-binding sites
GSU2504	9.4 up	*omcS*	cytochrome *c*, 6 heme-binding sites
GSU2501	6.3 up		cytochrome *c*, 6 heme-binding sites
GSU2584	5.3 up		lipoprotein, putative
GSU3214	5.3 up		cytochrome *c*, 3 heme-binding sites
GSU0710	5.1 up		conserved hypothetical protein
GSU0595.1	4.9 up		conserved hypothetical protein
GSU3410	4.6 up		conserved hypothetical protein
GSU0919	4.5 up		conserved hypothetical protein
GSU2586	4.0 up		hypothetical protein
GSU2497	3.8 up		lipoprotein, putative
GSU1018	3.8 up		hypothetical protein
GSU0594	3.7 up		cytochrome *c*, 7 heme-binding sites
GSU2811	3.6 up		cytochrome *c*, 2 heme-binding sites
GSU3409	3.6 up		conserved hypothetical protein
GSU2731	3.4 up	*omcC*	membrane-associated cytochrome *c*, 12 heme-binding sites
GSU1024	3.3 up	*ppcD*	cytochrome *c*, 3 heme-binding sites
GSU1947	3.2 up		hypothetical protein
GSU0193	3.2 up		L-sorbosone dehydrogenase, putative
GSU0767	3.0 up		putative porin
GSU2725	2.9 up		cytochrome *c*, 5 heme-binding sites
GSU2732	2.7 up		cytochrome *c*, 8 heme-binding sites
GSU1817	2.4 up		outer membrane lipoprotein, Slp family
GSU2743	2.4 up		cytochrome *c*, 1 heme-binding site
GSU2724	2.4 up		cytochrome *c*, 13-15 heme-binding sites
GSU1945	2.2 up		fibronectin type III domain protein
GSU2536	2.2 up		dienelactone hydrolase family protein
GSU2882	2.2 up	*omcG*	cytochrome *c*, 14-18 heme-binding sites
GSU0746	2.1 up		cytochrome p460, 1 heme-binding site
GSU0068	2.1 up		cytochrome *c*, 4 heme-binding sites
GSU2886.1	3.0 down		cytochrome *c*, 7 heme-binding sites
GSU2887	2.7 down		cytochrome *c*, 27 heme-binding sites
GSU2076	2.3 down	*omcZ*	cytochrome *c*, 7-8 heme-binding sites

Table 6. A short list of gene products predicted to reside in the ECM, for which transcripts were changed significantly in the microarray data (Δ*cheR5 vs.* WT). There are 49 genes in this table; hypothetical proteins with less than 3.0 fold changes are not included (see full list of 176 genes with significant (≥ 2-fold, $p \leq 0.005$) changes in expression in Appendix C.

Identification of chemoreceptors signaling through the Che5 pathway

In a Che pathway, MCPs sense signal molecules in the external environment or inside cells. In *E. coli*, MCPs are found co-localized with other chemotaxis signaling proteins at a specific position in the cell perimeter (Maddock & Shapiro, 1993b, Sourjik & Berg, 2000). In *E. coli*, all five MCPs form cluster with a single set of chemotaxis signaling proteins (Sourjik & Berg, 2000, Maddock & Shapiro, 1993a, Maddock & Shapiro, 1993b). In bacteria with more than one *che* cluster, and more than one class of MCPs (classified based on the length of MA domain), form distinct clusters in distinct subcellular locations (Guvener & Harwood, 2007, Guvener et al., 2006, Wadhams et al., 2003).

Typically, at least one *mcp* is found in a *che* cluster. In the *che5* cluster, however, there is no *mcp* identified. *G. sulfurreducens* contains 34 genes coding for MCPs, which are sorted into four major classes, including 8 MCPs in class 34H, 20 MCPs in class 40H, 4 MCPs in class 40+24H, and 1 MCP in class 44H (Alexander & Zhulin, 2007, Tran et al., 2008). We expect that only one class of MCPs would sense signal molecules for the Che5 pathway to regulate gene expression based on requirement for receptor interaction. We searched for the identity of this class by looking at genomes of bacteria that have *che5*-like clusters with identified *mcp* components. *Plesiocystis pacifica* was a species of special interest: its genome has only one *che* cluster, which is *che5*-like, and a single orphan *mcp* which codes for a 40+24H class MCP (NCBI identification number, gi_149923542). It is highly likely that this MCP is the receptor for the Che5-like signaling complex in *P. pacifica*. When this MCP sequence was used as a BLAST query against a database of all the proteins of *G. sulfurreducens*, the four MCPs with highest

scores are GSU2372, GSU0766, GSU0583 and GSU1704, which belong to one class of MCP (class 40+24H) (Tran et al., 2008), Figure 24. These MCPs were predicted to contain two transmembrane regions: GSU2372 and GSU0766 have a regular periplasmic domain (~ 200 aa), while GSU1704 and GSU0583 have smaller ones (10-20 aa). GSU1704 has a GAF domain in the cytoplasm (see the domain architecture of these proteins in Figure 25). Genes coding for these proteins were individually disrupted by replacement with antibiotic resistance cassettes. Strains with Δ*gsu1704* and Δ*gsu2372* showed phenotypes similar to those of Δ*cheR5*, Δ*cheW5b*, and Δ*cheA5*, in which cells are more aggregated and produce more OmcS and less OmcZ in the ECM (Figure 16 & 26). The Δ*gsu0766* also produced more OmcS and less OmcZ compared to the WT, but to a lesser extent than Δ*gsu1704* and Δ*gsu2372*, while the Δ*gsu0583* does did not significantly affect OmcS and OmcZ production (Figure 22 & 23). Similarly to the deletion mutants of Δ*cheA5*, Δ*cheW5b*, and Δ*cheR5*, deletion mutants of Δ*gsu1704* and Δ*gsu2372* increased transcription of *omcS* (Figure 22) and decreased transcription of *omcZ* (Figure 23).

	Score (bits)	E vaule
GSU0766| methyl-accepting chemotaxis protein	262	4e-71
GSU0583| methyl-accepting chemotaxis protein	261	6e-71
GSU2372| methyl-accepting chemotaxis protein	244	1e-65
GSU1704| GAF - sensor methyl-accepting	167	2e-42
GSU1029| methyl-accepting chemotaxis protein	142	6e-35
GSU1041| methyl-accepting chemotaxis protein	132	6e-32
GSU0935| methyl-accepting chemotaxis protein	130	3e-31
GSU0582| methyl-accepting chemotaxis protein	129	3e-31
GSU1374| methyl-accepting chemotaxis protein	129	4e-31
GSU1033.1| methyl-accepting chemotaxis protein	128	8e-31
GSU2579| methyl-accepting chemotaxis protein	127	1e-30
GSU0400| methyl-accepting chemotaxis protein	126	3e-30
GSU1033| methyl-accepting chemotaxis protein	125	4e-30
GSU0750| methyl-accepting chemotaxis protein	125	6e-30

Figure 24. A short list of results from BLAST of the only MCP of *Plesiocystis pacifica* against *G. sulfurreducens* proteins. Proteins in the red box were targeted for gene deletion in this study.

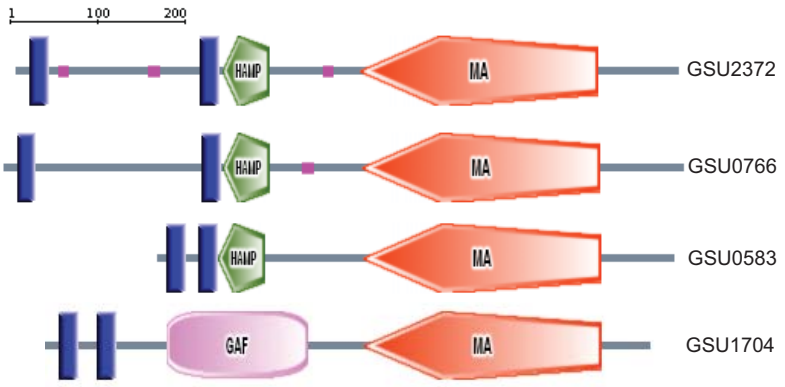


Figure 25. Domain architecture of MCP class 40+24H of *G. sulfurreducens*

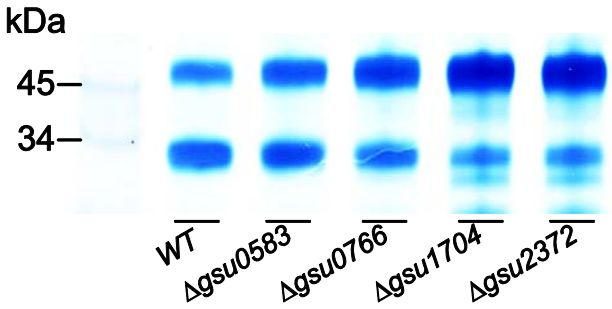


Figure 26. Heme-stained gel of ECM from *mcp* mutants. Similarly to the Δ*cheR5* strain, cells of Δ*gsu1704* and Δ*gsu2372* produce more OmcS and less OmcZ than the WT.

Discussion

The chemotaxis pathway that regulates bacterial movement toward chemical attractants or away from repellents is well-characterized, particularly in *E. coli*. Chemotaxis-like genes have been identified in most motile bacteria with a greater number and more complexity than the model species for chemotaxis studies, *E. coli*. Accumulated studies on species with multiple *che* clusters have revealed that cells utilize Che-like pathways to regulate diverse cellular functions. Besides a typical function in regulating chemotaxis, they are involved in biofilm formation, biogenesis of the motility apparatus, and gene regulation. Che-like pathways have been implicated in regulation of gene expression in bacterial species such as *M. xanthus*, *R. centenum*, and *Synechocystis* PCC6803. The Che3 cluster of *R. centenum* is involved in gene regulation for cyst cell development (Berleman & Bauer, 2005b); the Che3 and *dif* clusters of *M. xanthus* (Black et al., 2006, Kirby & Zusman, 2003, Yang et al., 2000) are involved in regulation of developmental genes and fibrils, respectively. The Tax1 and Tax2 clusters of *Synechocystis* PCC6803 (Bhaya et al., 2001, Chung *et al.*, 2001) have roles in regulating gene expression of type IV pili. However, in most cases, no functions parallel to those of Che proteins in the *E. coli* chemotaxis pathway have been shown in Che-like pathways that regulate gene expression.

G. sulfurreducens contains a total of ~ 70 *che*-like genes. The laboratory strain does not have flagella and is non-motile. When grown in acetate-fumarate liquid medium (at 30^{o}C), cells are planktonic, and evenly distributed in the culture tubes. To study the function of chemotaxis genes, individual *cheA* genes (*gsu0296*, *gsu1290*, *gsu2222*, *gsu3199*) were deleted by replacement with an antibiotic resistance cassette. Under

laboratory conditions, with the exception of the Δ*cheA5 gsu2222* which elevates cell aggregation, no apparent phenotype was observed in other Δ*cheA* strains. *Gsu2222* is co-transcribed with seven other *che* genes: *cheR5* (*gsu2215*), *cheB5* (*gsu2214*), *cheW5a* (*gsu2218*), *cheW5b* (*gsu2220*), *cheY5a* (*gsu2212*), *cheY5b* (*gsu2219*), *cheY5c* (*gsu2223*), together with 8 non-*che* genes in a cluster from *gsu2210* to *gsu2225* (Figure 14). In most *che* clusters, *mcp* genes are found associated with other *che* genes; in the *che5* cluster, however, no *mcp* is identified. *Che5*-like clusters were exclusively found in *δ-proteobacteria* (Tran et al., 2008); however, their function is not yet identified. We show here evidences that the Che5 proteins together with putatively one MCP class of *G. sulfurreducens* regulate gene expression of OmcS and OmcZ, which are two important outer membrane *c*-type cytochromes for bacteria to grow with an insoluble electron acceptor and to make a conductive biofilm. In addition, the Che5 pathway may also regulate a set of genes that make up extracellular materials and biofilm. Functional parallels of homologues in the *G. sulfurreducens* Che5 signaling pathway and the *E. coli* chemotaxis pathway are present.

In *E.coli*, the Che signaling pathway regulates the direction of cell movement by regulating the orientation of flagellar rotation. When there is no chemotaxis signal, *E.coli* performs random walking with alternation between smooth swimming (flagella rotate counter-clockwise), and tumbling (flagella rotate clockwise). Mutations that lead to inactivation of CheA, including deletions of *mcp* genes, *cheA*, *cheR*, *cheW*, and *cheY* result in cells with a smooth swimming bias. In contrast, mutants that increase the activity of CheA, including Δ*cheB* and Δ*cheZ*, result in cells with a tumble bias. We show that a similar logic is true for the Che5 pathway that regulates gene expression (Table 7).

Deletion of *mcp* genes, *cheA5*, *cheW* genes, and *cheR5* of the *che5* system leads to overexpression of OmcS and downregulation of OmcZ. Conversely, deletion of *cheB5* downregulates OmcS and upregulates OmcZ.

Overexpression of CheR in *E. coli* leads to cell tumbling, similar to the phenotype of the Δ*cheB* strain. We observed a similar pattern in the Che5 pathway, in which overexpression of CheR5 causes cells to produce more OmcZ and less OmcS than the WT, which is similar to the phenotype of a Δ*cheB5* deletion in *G. sulfurreducens* (Figure 21b).

In addition, microrarray data comparing Δ*cheR5* and WT strains show that there are 176 genes expressed at significantly different levels ($p \leq 0.005$, ≥ 2 fold changes) in Δ*cheR5* cells, listed in Appendix C. Forty-nine of them are predicted to reside in the extracellular space by the PSORTb program (Gardy et al., 2005). OmcS and OmcZ, which have been shown to be differentially regulated throughout the study, are among them. Furthermore, mutant strains of Δ*cheA5*, Δ*cheR5*, and Δ*cheW5b* overproduce filmentous materials, which were frequently observed in the biofilm of WT cells at elevated levels compared to planktonic cells (of WT). We hypothesize that the Che5 pathway regulates the expression of proteins involved in making up ECM and biofilm, and propose a model that could explain our data (Figure 27).

***E. coli* (motility)**		***G. sulfurreducens* (gene regulation)**	
Deletion mutant	Swim phenotype	Phenotype	Deletion mutant
Δ*mcp*	Smooth	**OmcS** up, **OmcZ** down	Δ*gsu1704; Δgsu2372*
Δ*cheR*	Smooth	**OmcS** up, **OmcZ** down	Δ*cheR5*
Δ*cheA*	Smooth	**OmcS** up, **OmcZ** down	Δ*cheA5*
cheW	Smooth	**OmcS** up, **OmcZ** down	Δ*cheW5a,* Δ*cheW5b*
Δ*cheY*	Smooth		
Δ*cheB*	Tumble	**OmcS** down, **OmcZ** up	Δ*cheB5*
Δ*cheZ*	Tumble	**OmcS** down, **OmcZ** up	Δ*cheY5c*
++*cheR*	Tumble	**OmcS** down, **OmcZ** up	++*cheR5*

Table 7. Summary of *G. sulfurreducens che5* mutant phenotypes with comparison to chemotaxis mutants in *E. coli*

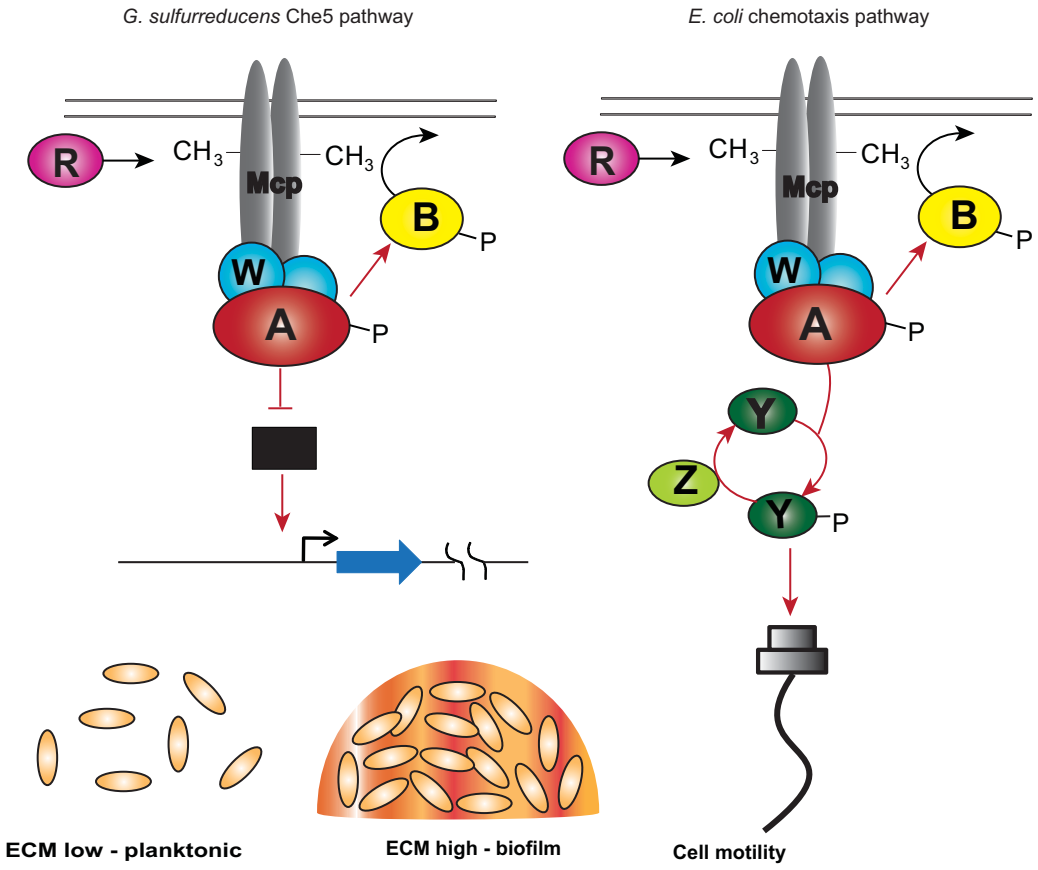


Figure 27. Proposed model for Che5 pathway regulation of gene expression, compared with the *E. coli* chemotaxis pathway

By analogy to the chemotaxis signaling pathway of *E. coli*, we expect that signals for the Che5 pathway of *G. sulfurreducens* are sensed via one class of MCPs (40+24H), that bind to CheA5 (GSU2222) via CheWs (CheW5a and CheW5b), which regulate activity of CheA5. The signal is then transduced to a cognate response regulator of CheA5 to regulate gene expression. The cells adapt to the different stimulus concentrations by modulating the level of covalent modification of MCPs, with contributions from CheR5 (GSU2215), CheB5 (GSU2214) and, indirectly, the CheYs (CheY5a, CheY5b, and CheY5c). How the Che5 pathway regulates gene expression is unclear, and the cognate response regulator of CheA5 is not yet identified. There are three *cheYs* in the *che5* cluster; if they were direct cognate response regulators of CheA5, and they targeted the genes that the Che5 pathway regulates, deletion mutants of *cheY5* genes should be phenotypically similar to a mutant of *cheA5* and whole cluster deletion mutants. However, we did not observe this with *cheY5* deletions (Figure 22). Phenotypes observed for the *cheY5* mutants, particularly Δ*cheY5c*, suggest their roles to be analogous to CheZ of *E. coli*. They could function as phosphate sinks similar to CheY1 of *S. meliloti* (Sourjik & Schmitt, 1998). Experiments are needed to test this hypothesis.

An outstanding question is, why does *G. sulfurreducens* use a Che-like pathway to regulate gene expression, and what is the advantage of using this pathway over a typical two-component system? As with *E. coli* chemotaxis, which allows the cells to be in a new state of movement transiently, the Che5 pathway could enable *G. sulfurreducens* to increase/decrease the expression of genes for a finite period of time when needed, and allow the cells to get back to basal level. This mechanism may also enable cells to colonize a surface to make a biofilm, or to leave established niches when necessary.

Previous studies have shown that bacteria can leave the biofilm to become planktonic (biofilm dispersal) depending on the availability of nutrients, and that planktonic cells and biofilm cells are different at the proteomic level (Sauer *et al.*, 2004), although the mechanism is not yet understood.

OmcS is an important outer membrane *c*-type cytochrome; it plays an essential role in transferring electron from inside the cell to insoluble electron acceptors, such as Fe(III) oxide and the anode of microbial fuel cells (MFCs) (Krushkal et al., 2008). When cells are grown with insoluble electron acceptors, they express OmcS (Mehta et al., 2005). Deletion of *omcS* inhibits growth on Fe(III) oxide and creates a defect in current production by a MCF (Mehta et al., 2005), (Holmes *et al.*, 2006). However, OmcS is not detected when soluble Fe(III) is used as electron acceptor, and deletion of *omcS* does not affect cells growing in this medium (Mehta et al., 2005). It appears that *G. sulfurreducens* is very efficient when it comes to expression of *omcS*; it only expresses this gene when needed. Probably the Che5 pathway is used for this purpose.

Strains with Δ*cheW5a* and Δ*cheW5b* deletions both exhibit effects on OmcS and OmcZ expression, but to different extents (Figure 22 & 23). The Δ*cheW5b* mutant increased *omcS* transcription 18-fold, in contrast to a less than 2-fold increase in the Δ*cheW5a* mutant. Transcription of *omcZ* decreased 3-fold in the Δ*cheW5b* mutant, and less than 2-fold in the Δ*cheW5a* mutant.These data indicate that the presence of two *cheW* genes in the cluster is not redundant; they have different roles in regulating genes. The *cheOp2* cluster of *R. sphaerodes* also contains two *cheW* genes, and a study has shown that in different environments, the *cheW* genes contribute differently to regulation of chemotaxis (Martin et al., 2001).

There are 8 non-*che* genes in the *che5* operon; products of these genes contain conserved domains found in other species, but their functions are not well understood. GSU2210 is a *c*-type cytochrome protein; GSU2213 contains a GAF (cyclic GMP, adenylyl cyclase, FhlA) domain (Hurley, 2003); GSU2216 is comprised a HEAT-like repeats; GSU2217 contains a response receiver domain; GSU2221 belongs to an ATPase family. GSU2224 also contains a receiver domain, in addition to the PATAN domain, which is predicted to be involved in signal transduction machinery, and the FRGAF domain, named after FrgA of *M. xanthus*, a GAF domain-containing protein involved in regulation of fruiting body formation (Makarova et al., 2006). GSU2225 belongs to the GTPase EngA sub-family, which is found in both eukaryotes and prokaryotes. In bacteria, its function is not yet elucidated. Studies of the *E. coli* EngA homolog, Der (Robinson *et al.*, 2002), and *Neisseria gonorrhoeae* EngA (Mehr *et al.*, 2000) showed that they are essential for cell viability. It is predicted that GTPase EngA could be involved in ribosome assembly or stability (Caldon & March, 2003). The creation of a *gsu2225* deletion mutant in *G. sulfurreducens*, and the absence of any effect of the mutation on growth in standard culture medium, desmonstrate that it is not an essential gene for cell viability, so it is likely to have a different role than those reported in *E. coli* and *N. gonorrhoeae*.

Deletion mutants of non-*che* genes in the *che5* cluster have no visible phenotype when growing in liquid medium (acetate-fumarate, at 30^{o}C), and show no significant effect on OmcS and OmcZ expression, with the exception of Δ*gsu2216*. ECM of Δ*gsu2216* show a similar phenotype to the Δ*cheB* mutant, suggesting that GSU2216 could coordinate with Che5 to regulate gene expression of ECM. There is no other

homolog of *gsu2216* in the genome of *G. sulfurreducens*. GSU2216 contains 12 repeats of a HEAT-like domain (Figure 15). The HEAT repeat is found in subunits of cyanobacterial phycocyanin lyase, and other proteins and is predicted to be involved in mediating protein-protein interaction and assembly of multiprotein complexes (http://pfam.sanger.ac.uk/family?id=HEAT_PBS). How GSU2216 works with other Che proteins of Che5 is unknown and warrants further investigation. It is worth noting that GSU2216 homologs are present in all Che5-like clusters identified in *δ-proteobacteria*, and in all cases, located next to the *cheR* (Tran et al., 2008). It is likely that the function of GSU2216 in *G. sulfurreducens* is also present in other Che5-like pathways. In *M. xanthus*, beside the *che5*-like cluster, cluster 7 (Mxan_6958-6966) also contains a *gsu2216*-like gene (Mxan_6961), located next to *cheR* (http://img.jgi.doe.gov).

There are 34 *mcp* homologs in the genome of *G. sulfurreducens*, but none of them is found in the *che5* operon. We identified a class of MCPs (40+24H), including GSU0583, GSU0766, GSU1704 and GSU2372, that could couple with the Che5 complex to regulate gene expression. The phenotypes of Δ*gsu1704* and Δ*gsu2372* strains are similar to those of Δ*cheR5*, Δ*cheA5*, and Δ*cheW5a*, and Δ*cheW5b* strains, suggesting that GSU1704 and GSU2372 are the major sensors for the Che5 pathway, and that they both are essential for the formation of a signaling complex (MCPs, CheWs and CheA), upon which the signal is transduced from sensors to the cognate response regulator. The absence of one MCP leads to disruption of the signaling pathway. In this respect it seems that the Che5 pathway in *G. sufurreducens* is different from the chemotaxis pathway of *E. coli*. In *E. coli* and *S. typhimurium* there are 5 MCPs; deletion of one of the two major MCPs (Δ*tsr* or Δ*tar*) does not affect the signaling complex of the remaining MCPs with

CheW and CheA, and therefore the cells can still respond to the signals sensed by the remaining MCPs (Imae *et al.*, 1987). Strains with Δ*gsu0766* and Δ*gsu0583* have no significant impact on gene regulation of *omcS* and *omcZ*, but this may not mean that they are not signaling through the Che5 complex. It is probable that their absence does not affect formation of the Che5 complex with other MCPs in the same class, and thus does not impact the signaling pathway, as is the case for MCPs in *E. coli*.

The four MCPs predicted to couple with the Che5 complex have different n-terminal sensing domain architectures (Figure 25), suggesting that the input signals are diverse. GSU2372 and GSU0766 have a periplasmic region of an MCP (~200 aa), suggesting that they could sense signals from the external stimuli in the environment. GSU0583 has a small periplasmic domain that may sense signals in the periplasm or sense signals via other proteins binding to it. GSU1704 has a GAF domain in the cytoplasm, suggesting that it could sense small molecules such cGMP, cAMP, or c-di-GMP inside the cell.

A previous study suggested a function for an MCP (RppA-receptor for polysaccharide production) of *M. xanthus*: involvement in regulation of biosynthesis/assembly of polysaccharide (Kimura *et al.*, 2004). RppA also belongs to the 40+24H class of MCPs, as do MCPs of the Che5 pathway in *G. sulfurreducens*. However, how it regulates polysaccharide production and through which of the eight Che *M. xanthus* clusters of this MCP signals, has not yet been determined.

CHAPTER 4

PRELIMINARY DATA FROM CHEMOTAXIS STUDIES OF A MOTILE GEOBACTER SULFURREDUCENS STRAIN

Introduction

The ability of bacteria to sense chemical gradients and swim toward the favorable environment and away from the unfavorable one, which is known as chemotaxis, helps them to navigate to their niches that are optimal for their growth and survival. Previous studies showed that bacteria are chemotactic to various environmental pollutants such as polycyclic aromatic hydrocarbons, nitroaromatic compounds, petroleum-associated hydrocarbons, explosives and their respective metabolic intermediates/transformants (Bhushan *et al.*, 2004, Gordillo *et al.*, 2007, Grimm & Harwood, 1997, Samanta *et al.*, 2000). The ability of bacteria to swim toward a contaminant attractant via chemotaxis will increase bioavailability, and therefore enhance the biodegradation process. Indeed, chemotaxis has been considered to play an important role for indigenous bacteria to enhance *in situ* remediation, and has become a component in a bioremediation model (Singh, 2008).

Geobacter species have potential to bioremediate organic compounds and metal compounds. They are the dominate species in uranium bioremediation sites (Anderson et al., 2003). The genomes of *Geobacter* species contain genes encoding a motility apparatus (flagella and pili) and a large number of chemtoaxis genes, suggesting that these genes could play important roles in sustaining cell growth and survival in the changing environment. In fact, *G. metallireducens* and *G. uraniireducens* have flagella, and are motile (Shelobolina et al., 2008, Childers *et al.*, 2002). A previous study indicated

that *G. metallireducens* is chemotactic toward its spent terminal electron acceptor (Fe^{2+}) (Childers et al., 2002). Although it is not a direct response to the electron acceptor (Fe^{3+}), it was argued that the attraction to Fe^{2+}, a product of the metabolic process, could enhance bacterial access to the site of remediation (Childers et al., 2002). Perhaps the presence of a large number of chemotaxis genes enables *Geobacter* species to sense various environmental signals, and it could be one of the reasons why *Geobacter* species outcompete other bacterial species during *in situ* uranium bioremediation. To date, there no is information about which genes are involved in chemotaxis of *Geobacter*.

Geobacter species have been studied extensively not only because of their application in bioremediation, but for their potential to generate electricity in microbial fuel cells. Among them, *G. sulfurreducens* is most frequently studied. It is among the most productive pure cultures to produce electricity in MFCs (Nevin *et al.*, 2008), partly due to its ability to transfer electrons directly to the electrode, and due to the formation of a conductive biofilm (Reguera et al., 2005), and N. Malvankar unpulished data). In a recent study, during the selection process to improve power production by *G. sulfurreducens*, a new strain was isolated and designated KN400 (Yi *et al.*). One significant difference between KN400 and the original *G. sulfurreducens* DL1 strain is that KN400 produces flagella and is motile, while DL1 has no flagella and is non-motile (Yi et al., Caccavo et al., 1994b). KN400 is much more productive than DL1, in terms of electricity generation (Yi et al.). KN400 cells start to produce current and reach the maximum current in a much shorter period of time than DL1 (Yi et al.), suggesting that KN400 may have mechanisms that make the biofilm cover the electrode more quickly than DL1.

Biofilm formation is essential for current production in MFCs of *G. sulfurreducens*. The biofilm formation processes are well-studied in other bacteria (O'Toole et al., 2000). There are three major steps in the formation of biofilm: (1) the initial attachment, which forms a monolayer; (2) the formation of microcolonies; and (3) biofilm maturation. Flagellar motility has been found to be essential for the initial attachment, and also during biofilm maturation (O'Toole et al., 2000). Previous studies also demonstrated that chemotaxis plays important roles in biofilm formation (Kirov *et al.*, 2004, Merritt *et al.*, 2007). These findings, together with the observations in KN400 and DL1, suggest that flagella and chemotaxis could be part of the reason that KN400 is a much better electricigen than DL1.

Knowledge about chemotaxis in *Geobacter* can provide a better understanding about the physiological functions of chemotaxis genes and their contributions to bioremediation and electricity generation, and provide information toward optimizing conditions for more effective applications. Our chemotaxis study in *Geobacter* is based on *G. sulfurreducens* KN400, taking advantage of the availability of a genetic system for *G. sulfurreducens* (Coppi et al., 2001). Results of the study can extended to other *Geobacter* species, because of the high similarity in gene order and gene sequence identity among *Geobacter* species.

Methods

Strains and growth conditions

The plasmids and *G. sulfurreducens* strains used in this study are listed in Table 8. *G. sulfurreducens* strains were routinely cultured in acetate-fumarate medium at 30°C under strict anaerobic conditions as previously described (Lovley & Phillips, 1988a, Coppi et al., 2001).

Designation	Genotype or description	Source or reference
Strains		
DL1	*G. sulfurreducens*, non-motile	(Caccavo et al., 1994b)
pilA2	DL1, *gsu1496::kan*	Al-Challah unpublished data
KN400	*G. sulfurreducens*, motile	(Yi et al.)
Δ*flgE*	KN400, *gsu0419::kan*	This study
Δ*motA*	KN400, *gsu3027::kan*	This study
Δ*pilA*	KN400, *gsu1496::kan*	This study
Δ*gsu0296*	KN400, *gsu0296::kan*	This study
Δ*gsu1290*	KN400, *gsu1290::kan*	This study
Δ*gsu2222*	KN400, *gsu2222::kan*	This study
Δ*gsu3199*	KN400, *gsu3199::kan*	This study
Plasmids		
pCD341	Source of kanamycin resistance cassette; Kan^r	(Morales et al., 1991)

Table 8. Bacterial strains and plasmids used in chapter 4

Construction of *G. sulfurreducens* strains

G. sulfurreducens genomic DNA was purified using the MasterPure Complete DNA purification kit (Epicentre Technologies, Madison, WI). Mini Plasmid purification kits, PCR purification kits, and Qiaquick gel extraction kits (QIAGEN Inc. Valencia, CA) were used for plasmid DNA purification, PCR product purification, and gel extraction, respectively. Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs, Inc. (Beverly, MA). Primers were purchased from Sigma-Genosys (Sigma-Aldrich Co., St. Louis, MO) and Operon Biotechnologies, Inc. (Huntsville, AL). Taq DNA polymerase (QIAGEN Inc., Valencia, CA), and Phusion polymerase (Finnzymes) were used for PCR amplifications. The following conditions were used: 95°C for 5 min, followed by 30 cycles of 95°C for 30 s, 58°C for 30 s, and 72°C for a period of time (estimated based on the rate of ~ 1kb/min) and a final extension at 72°C for 10 min (for Taq DNA polymerase); and 98°C for 30 s, followed by 30 cycles of 98°C for 30 s, 58°C for 30 s, and 72°C for a period of time (estimated based on the rate of 1.5-2 kb/min), for Phusion polymerase.

Genes encoding homologues of FlgE (GSU0419), MotA (GSU3027), and 4 CheAs (GSU0296, GSU1290, GSU2222, and GSU3199) were individually disrupted by an antibiotic resistance cassette. A linear mutagenic fragment was generated and a single-step gene replacement was performed as previously described (Coppi et al., 2001), with some modifications. Briefly, the sequences of upstream and downstream regions and the antibiotic resistance cassette were amplified with the corresponding primers listed in Table 9, as demonstrated in the schematic (Figure 28). PCR products of the upstream region (with primers 1 & 2), antibiotic resistance cassette (with primers 3 & 4), and

downstream region (with primers 5 & 6) were digested with restriction enzymes, and then the three fragments were ligated. Ligation products were run on a DNA gel to extract a band at the right size (~2.1 kb). It was then amplified with the distal primers (1 and 6), gel-purified, and concentrated to a final concentration of 500 – 1,000 ng/μL DNA. Mutagenic fragment for generating *pilA* (*gsu1496*) deletion strain was prepared from strain pilA2 (Al-Chalah, unpublished data), which is an inframe deletion of pilA. Primers rlc45FpilR and rlc50R1497 (Al-Challah, unpublished data) were used to amplified mutagenic fragment; DNA of pilA2 strain was used as a template. PCR products were then gel-purified for a ~ 3.9 kb, and concentrate to a final concentration of 500 – 1,000 ng/μL. A total of ~ 1 μg of mutagenic fragment was transformed by electroporation into *G. sulfurreducens* KN400 as previously described (Coppi et al., 2001). After a recovery period, cells were plated on NBAF agar with the antibiotic concentrations described earlier (Coppi et al., 2001). Mutants were confirmed by PCR.

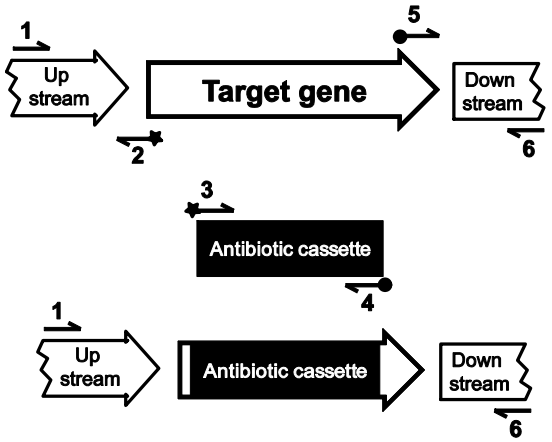


Figure 28. Schematic for making linear mutagenic fragments. Primers were designed for PCR of ~500 bp upstream of target genes (primers 1 & 2); ~500 bp downstream of target genes (primers 5 & 6) and the antibiotic resistance cassette (primers 3 & 4). A restriction site that is not found in the sequences of the upstream region and antibiotic resistance cassette was added to primers 2 and 3. Another restriction site that is not present in the sequences of the antibiotic resistance cassette and downstream region was added to primers 4 and 5.

Purpose	Primers and sequences	Descriptions
gsu0296::kan	HT55F 5'-GCTGGTGAGTTTCAAGCTAGAGGAAGAGG-3' HT55R 5'-CGATAAGCTTAATCGCCATGTGCTGCTCC-3'	To amplify GSU0296 upstream
	HT42F 5'-GCTGCAAGCTTTGGGATGAATGTCAGCTAC-3' HT22R 5'-GCACTCTAGAGCGGTGGAATCGAAATCTCG-3'	To amplify Kan
	HT56F 5'-GCTATCTAGACCTGAGCGGACAACGAGC-3' HT56R 5'-GCTGATGTCCGTACCGACGATTTCGATG-3'	To amplify GSU0296 downstream
gsu1290::kan	HT57F 5'-CGTCGATATTACCGGCCTGCAACTGC-3' HT57R 5'-GCATAAGCTTGTCCATTAGCCGAGCACC-3'	To amplify GSU1290 upstream
	HT42F and HT22R as above	To amplify Kan
	HT58F 5'-GCAGTCTAGACCGGTAGAATGTGCTCTGC-3' HT58R 5'-GGTAGATCAGGTTGCATCTGGGGAGAACC-3'	To amplify GSU1290 downstream
gsu2222::kan	HT61F 5'-CCATCCTCTACGGCGACAATCTGC-3' HT61R 5'-CGTAGTCGACGTGTTGGTCATGGAACC-3'	To amplify GSU2222 upstream
	HT22F 5'-GCTAGTCGACTGGGATGAATGTCAGCTAC-3' HT22R as above	To amplify Kan
	HT62F' 5'-GCTATCTAGAGGCAGCAGGACATCGTCATC-3' HT62R' 5'-GCATCAGGTAGAGCGTTTCCGTGAG-3'	To amplify GSU2222 downstream
gsu3199::kan	HT59F 5'-CGAGTGAACATCCGCGTTTCGAGGGATAC-3' HT59R 5'-GCTCAAGCTTGGACATGTCCATGTCTAGC-3'	To amplify GSU3199 upstream
	HT42F and HT22R as above	To amplify Kan
	HT60F' 5'-GCAGCTCTAGAGCAGAGGTATTCACCAAGC-3' HT60R' 5'-CGAGATCTCTAGCTCTGTGAGGAGGGACC-3'	To amplify GSU3199 downstream
gsu0419::kan	HT270F 5'-CGTGGTGAGTGTCAGAAACCGGACAGTC-3' HT270R 5'-GAATGTCGACGACGTCCGAGAAGAGCATCC-3'	To amplify GSU0419 upstream
	HT22F 5'-GCTAGTCGACTGGGATGAATGTCAGCTAC-3' HT22R as above	To amplify Kan
	HT271F 5'-GCTATCTAGACCTCTTCGAGGAGACCCTCG-3' HT271R 5'-CGAAGAAACGTACGAGAGGCGAGACAC-3'	To amplify GSU0419 downstream
gsu3027::kan	HT261F 5'-GCACTGCTGTAGCTAAAGTTTTCCGGTCC-3' HT272R 5'-GAATGTCGACTCAGGTGCAAACCTTCCAGC-3'	To amplify GSU3027 upstream
	HT22F and HT22R as above	To amplify Kan
	HT273F 5'-GCTGTCTAGACTTCGGCACAAAGATCAAGC-3' HT263R 5'-CATGATCCGCATGTCGGACGAATCCAGTACTG-3'	To amplify GSU3027 downstream
To distinguish between DL1 and KN400	Primer in contig 95 95-1F 2058540-1941 (1402) 95-2R 794-1735 (942)	To amplify a gene that is present in KN400 but not in DL1 (from H. Yi)

Table 9. List of primers used in chapter 4

Swarm plate assays for chemotaxis, and selection for motile cells

G. sulfurreducens KN400 stock cells were grown to late log phase in NBAF medium. Concentrated cells were prepared in anaerobic conditions by centrifugation with a VWR Galaxy mini centrifuge. Approximately 90% volume of supernatant was removed and cells were re-suspended in the remaining medium. 7 uL of concentrated cells were stabbed with a 10 μL pipet tip into the middle of a Petri dish containing NBAF with 0.3% (wt/vol) Difco^{TM} agar noble (Becton, Dickinson and Company, MD, US), which was poured and left overnight to solidify. Swarm plates were incubated at 30^{o}C and kept in an anaerobic glove bag until a clear ring was formed (after about 7-10 days). To select motile cells for further study, cells in the outer ring of the swarm plate were picked up with a blunt pipette tip, grown in the liquid medium, and purified by repeating the swarm plate assay a second time. This time, cells from the outer ring were streaked on an NBAF 1.5% agar plate for single colony selection. One colony was selected and re-streaked to an agar plate for colony purification. Three colonies were then picked up, characterized by looking at single cells by TEM (as described in Chapter 3), and tested on swarm plates. One isolate was used throughout the study.

Results and discussion

KN400 forms chemotactic rings on the swarm plates

To test whether KN400 cells behave chemotactically, semi-solid agar plates were used. As seen in Figure 29, cells form clear rings after several days of incubation. This is typical of *E. coli* chemotactic behavior when responding to an increasing gradient of chemo-attractants. The rings formed are due to the cell movement toward a gradient of nutrients in the plate. As the cells grow, nutrients are depleted, and therefore the cells spread out to search for food. These data suggest that KN400 is chemotactic, although the specific attractants are unknown. It is predicted that KN400 cells may respond to acetate, fumarate or salt compounds, minerals or vitamins in the medium.

Figure 29. Swarm assay of strain KN400. An arrow indicates the ring formed by chemotactic processes.

Observed chemotactic behavior is due to flagellum-based motility

Previous studies indicated that chemotactic motility is driven either by flagella or by type IV pili (Zusman et al., 2007, Szurmant & Ordal, 2004). The observation that KN400 cells were able to swarm out on the soft agar plates could be due to either flagellar or pilus-based motility. To identify the motility apparatus that leads to chemotactic behavior in *G. sulfurreducens* KN400, we generated deletion mutants of type IV pilin, which is encoded by *gsu1496* (*pilA*), and of flagellar genes (*gsu0419* and *gsu3027*). The flagellar proteins that were targeted in this study are shown in the schematic.

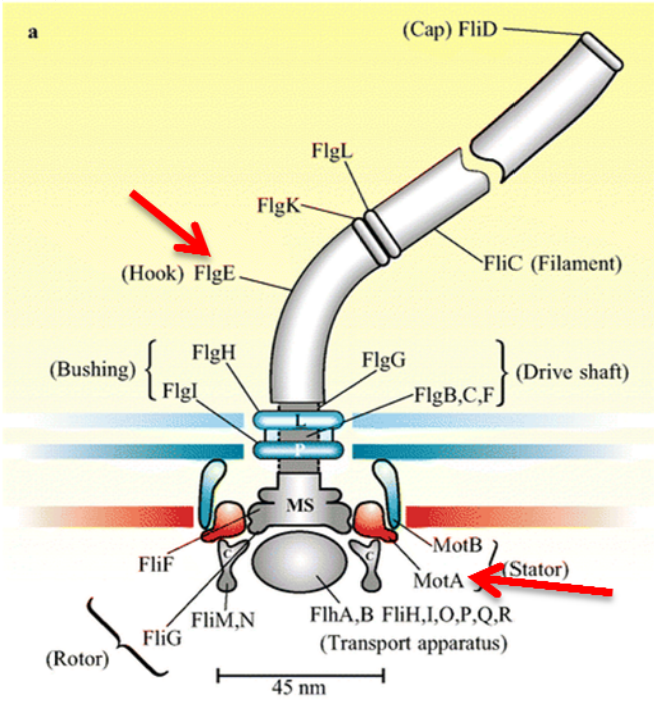


Figure 30. Schematic illustrating the protein components of the flagellar apparatus of *E. coli* and *S. enterica* serovar Typhimurium (Berg, 2000). The red arrows point to the targeted genes in this study.

To examine the success of generating the deletion mutants, we used PCR with distal primers (primers 1 and 6 in Figure 28) with genomic DNA from WT and mutants as templates. For evaluating Δ*pilA*, primers (LA43F and LA50R and rLC47R and rLC50R from Al-Challah L.) were used. PCR products of WT and mutants are predicted to be 1.5 kb and 3.9 kb respectively. Primers HT270F and HT271R that generate PCR products of 1.8 kb in WT and 2.1 kb in mutant were used to evaluate the Δ*flgE* strain. To check the Δ*motA* strain, HT261F & HT263R were used, which generate PCR products of 1.4 kb for WT and 2.1 kb for the mutant. Figure 31 shows the expected bands of KN400 and mutants, confirming the success of generating deletion mutants of *pilA*, GSU0419 and GSU3027. In addition, to confirm that mutants were generated in the KN400 background, we used primers (CP95-1F & CP95-1R) that amplify a gene of ~1.5 kb that is present in KN400 but not in the genomic strain DL1. As seen in Figure 31, the ~1.5 kb band is detected in all mutant strains, which demonstrates that they are of KN400 origin, but not DL1 background.

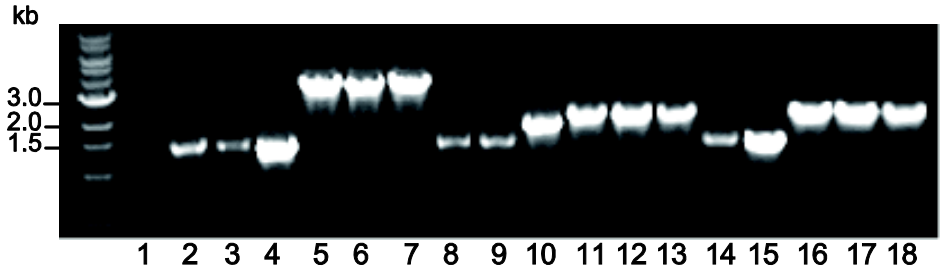


Figure 31. DNA gel resolving PCR products from WT and mutants. The standard marker is on the left. The table below details the gel loading order (left panel) and predicted sizes of the PCR products (right panel).

Gel loading order	**Predicted size**
Lane 1. DL1 with primers CP95-1F & CP95-1R	None
Lane 2. KN400 with primers CP95-1F & CP95-1R	1.5 kb
Lane 3. KN400 with primers (LA43F & LA50R)	1.5 kb
Lane 4. KN400 Δ*pilA* (colony 1) with primers CP95-1F & CP95-1R	1.5 kb
Lane 5, 6, and 7. KN400 Δ*pilA* colonies 1, 2, and 3 with primers (LA43F & LA50R)	3.9 kb
Lane 8, and 9. KN400 Δ*flgE* colonies 1 and 2 with primers CP95-1F & CP95-1R	1.5 kb
Lane 10. KN400 with primers HT270F & HT271R as control	1.8 kb
Lane 11, 12, and 13. KN400 Δ*flgE* colonies 1, 2 and 3 with with primers HT270F & HT271R	2.1 kb
Lane 14. KN400 Δ*motA* colonies 1 with primers CP95-1F & CP95-1R	1.5 kb
Lane 15. KN400 with primers HT261F & HT263R as a control	1.4 kb
Lane 16, 17, and 18. KN400 Δ*motA* colonies 1, 2 and 3 with primers HT261F & HT263R	2.1 kb

GSU0419 encodes a homolog of FlgE, a protein component of the flagellar hook. Deletion of *flgE* in *E. coli* and *S. typhimurium* produces cells without flagella (Macnab, 1996, Berg, 2000). GSU3027 codes for a homolog of MotA, one of the two motor proteins that play an essential role in rotating flagella in *E. coli* and *S. typhimurium*. Deletion of *motA* produces paralyzed cells, i.e., cells have flagella but are nonmotile (Macnab, 1996). Deletion of *pilA*, GSU0419, and GSU3027 does not affect the growth rate in NBAF medium (with comparison to the WT), under the conditions tested (data not shown). Observed by TEM, Δ*pilA* and Δ*motA* cells have flagella; however, no flagella were observed for cells of the ΔflgE strain (Figure 32).

To test whether these mutants affect chemotactic behavior, swarm assays were employed. As seen in Figure 33, deletion of *pilA* does not affect the formation of chemotactic rings, while deletion of *flgE* or *motA* abolishes the ability of KN400 to swarm out. Our data suggest that the swarming out of KN400 is due to flagellar motility, but not due to twitching motility dependent on type IV pilin.

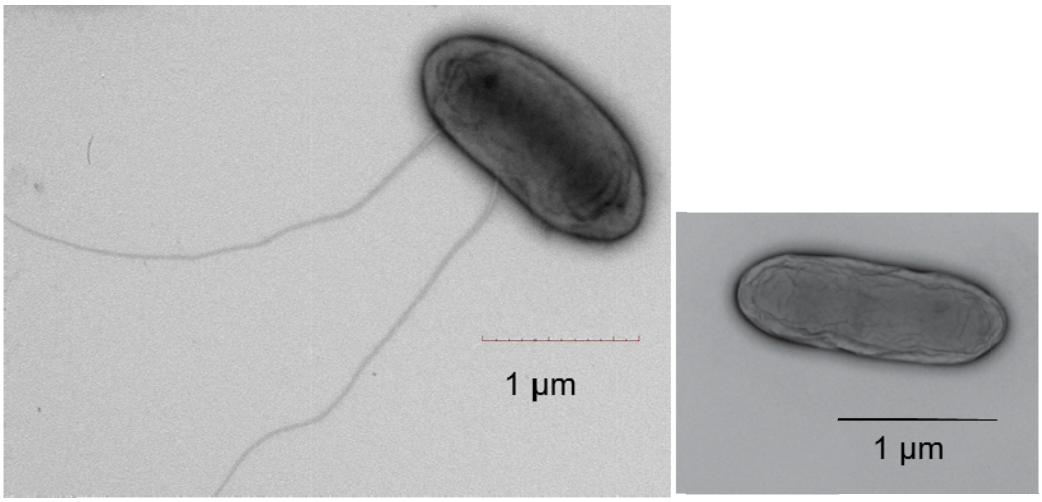


Figure 32. Representative TEM images of KN400. Δ*pilA* and Δ*motA* cells show the presence of flagella (left), and their absence from Δ*flgE* cells (right).

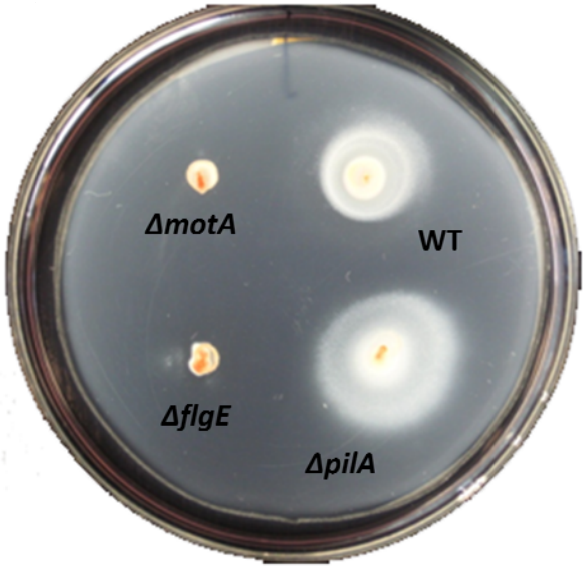


Figure 33. Swarm assay of KN400 WT, Δ*pilA*, Δ*flgE*, and Δ*motA* strains. This picture was taken after 9 days of incubation

Che1 may be involved in regulating chemotaxis in *G. sulfurreducens* KN400

There are five major *che* clusters in *G. sulfurreducens*, with 4 *cheA* genes located in 4 different clusters, predicted to regulate four different Che-like pathways. To identify which gene cluster is involved in regulating chemotactic motility, deletions were made in individual *cheA* genes, the central components of each pathway. Primers for generating *cheA* mutants are listed in Table 9. Figure 34 demonstrates the success of mutant generation.

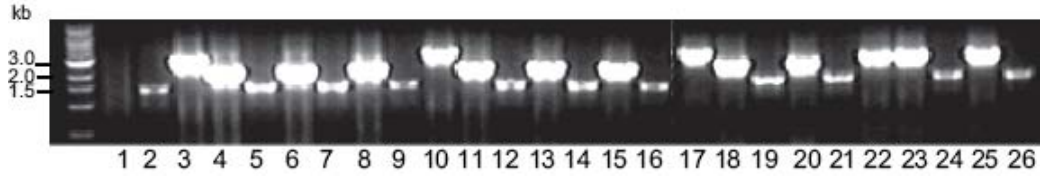


Figure 34. PCR products from WT and *cheA* mutants. The standard marker is on the left. Gel loading order and predicted sizes are below.

Gel loading order	**Predicted size**
Lane 1. DL1 with primers CP95-1F & CP95-1R (negative control for KN400)	None
Lane 2. KN400 with primers CP95-1F & CP95-1R (control for KN400)	1.5 kb
Lane 3. KN400 with primers HT55F & HT56R (as control)	2.8 kb
Lane (4, 5); (6,7); (8,9). KN400 Δ*gsu0296* colonies 1, 2 and 3 with primers HT55F & HT56R; and CP95-1F & CP95-1R	(2.1; 1.5 kb)
Lane 10. KN400 with primers HT57F & HT59R (as control)	3.1 kb
Lane (11,12); (13,14); (15,16). KN400 Δ*gsu1290* colonies 1, 2, and 3 with primers HT57F & HT58R; and CP95-1F & CP95-1R.	(2.1; 1.5 kb)
Lane 17. KN400 with primers HT61F & HT62R' (as control)	3.1 kb
Lane (18,19); (20,21). KN400 Δ*gsu2222* colonies 1 and 2 with primers HT61F & HT62R', and CP95-1F & CP95-1R	(2.1; 1.5 kb)
Lane 22. KN400 with primers HT59F & HT60R' (as control)	2.5 kb
Lane (23,24); (25,26). KN400 Δ*gsu3199* colonies 1 and 2 with primers HT59F & HT60R', and CP95-1F & CP95-1R	2.2 kb

Because the gel showed no difference between the PCR bands from KN400 WT and the Δ*gsu3199* strain, further experiments were carried out. Primers 1 and 4 (as shown in the schematic, corresponding to HT59F and HT22R for the case of Δ*gsu3199*) were used to amplify genomic DNA from WT and Δ*gsu3199*. The PCR product should only be detected in the mutant but not in the WT, and our result showed the expected band (data not shown). This suggests that *gsu3199* was successfully deleted and replaced by a kanamycin cassette as designed.

By TEM, flagella are observed in all Δ*cheA* strains. There is no significant growth rate difference seen in mutants compared to the WT. To test whether the mutations affect cell motility, stationary phase cells were harvested and stabbed into swarm plates. As seen in Figure 35, deletion of *gsu0296* abolished cell swarming, while other mutations did not affect cell motility under the conditions tested. This result suggested that GSU0296, together with other genes in the *che1* cluster, regulates chemotaxis of *G. sulfurreducens*. It is worth noting that to date, the *che1* cluster has not been found in any species other than *Geobacteraceae* (Tran et al., 2008). To our knowledge it is the first time a *che* cluster involved in regulation of chemotaxis in *Geobacter* species is identified. More information is needed toward better understanding of the detailed mechanism of the Che1 pathway. The cluster encodes, in addition to a CheA, homologs of CheR, CheB, and CheW, and a non-Che protein, which contain an HD domain (Tran et al., 2008). Like the *che5* cluster described in Chapter 3, thre are no genes for MCPs in the *che1* cluster.

Figure 35. Swarming assay of KN400 and *cheA* mutants. Cells were spotted clockwise: WT (1); Δ*gsu0296* (2); Δ*gsu1290* (3); Δ*gsu2222* (4), and Δ*gsu3199* (5). This image was taken after 10 days incubation.

Conclusion

We showed that *G. sulfurreducens* KN400 is chemotactic, and this behavior is flagellum-dependent. Our preliminary data indicate that *Geobacter* may use its unique *che* cluster, *che1*, to regulate chemotaxis. GSU1290, which is in an *E. coli*-like cluster, does not show any effect on chemotaxis assays under the conditions tested in this study, but that does not preclude the possibility that it is involved in chemtoaxis. It may function in other environmental conditions (not yet tested in this study), as has been found in other bacteria such as *R. sphaeroides* (Harrison *et al.*, 1999, Wadhams et al., 2003).

The ability of *Geobacter* to swim along chemical gradients, and the presence of a large number of chemoreceptors that might enable them to sense various signal molecules, could explain why *Geobacter* species outcompete other species at the site of *in situ* uranium bioremediation. Chemotaxis clearly should be taken into consideration in modeling *in situ* bioremediation. It could also be applied to enhance bioremediation processes by generating a chemoattractant gradient to help guide bacteria to the site of bioremediation.

Previous studies demonstrated that in MFC mode, *G. sulfurreducens* KN400 quickly produces an electrical current and at significantly higher current density than DL1 (Yi et al.). It is not clear why KN400 is more productive than DL1 in terms of electricity generation. Chemotaxis has been shown to play an important role in the formation of biofilm in some bacteria (Kirov et al., 2004, Merritt et al., 2007). The ability of *G. sulfurreducens* KN400 to swim with direction may help it to find the electrode quickly to establish the initial attachment for the development of biofilm, which is required for *G. sulfurreducens* to make current; this could be one of the reasons why *G.*

sulfurreducens KN400 begins to produce current in a shorter period of time than the non-motile DL1 strain. More detailed studies of how the functions of chemotaxis proteins contribute to bioremediation and electricity generation are warranted.

CHAPTER 5

FUTURE STUDIES

Examination of the formation of MCP class-specific clusters

Our study (reported in Chapter 2) showed that *Geobacter* species have a large number of genes encoding MCPs, and based on the length of cytoplasmic MA domains, at least 4 MCP classes were identified in each *Geobacter* species. Protein complex formation of MCP and CheA via CheW is essential for signal generation, amplification and adaptation in the chemotaxis pathway. In *E. coli*, all five MCPs of the same class form a complex with other chemotaxis signaling proteins, and are located at the pole of the cell to regulate the direction of its movement. Previous studies demonstrated that in bacteria with multiple *che* clusters, their functions are diverse, and they tend to segregate into different locations within the cell (Wadhams et al., 2003, Guvener et al., 2006). This suggests that the cells may have mechanisms, yet to be identified, to avoid cross-talk among Che-like pathways. We speculated that MCPs of the same class, together with their signaling protein partners, are likely to form a complex, segregated in a specific location within the cell. This could be a mechanism that cells use to avoid cross-talk. Knowledge about factors that lead to the specific segregation of Che proteins could contribute to an understanding of the functioning of multiple *che* clusters in bacteria. Our hypothesis could be tested by using single or combined techniques described below:

1. Identification of the locations of MCPs in *Geobacter* species.
    - Immunogold electron microscopy has been used to study the localization of MCPs in bacteria such as *E. coli* and *R. sphaeroides* (Alley *et al.*, 1992, Harrison et al., 1999). Previous studies demonstrated the success of using an antibody

which was raised against a highly conserved domain of *E. coli* Tsr for recognition of MCPs in *R. sphaeroides*, as evidenced by western blot and immunogold labeling. It is very likely that *Geobacter* MCPs could react with the *E. coli* Tsr antibody (Harrison et al., 1999), and immunogold electron microscopy could be used to identify the locations of MCPs in *Geobacter*. We predict that four or more clusters of gold particles will be found in different locations within *Geobacter* cells, corresponding to the number of MCP classes and *che* clusters.

- Jensen's group developed a method to enable visualization of the topology of MCP complexes in *Caulobacter crescentus* intact cells, using cryo-electron microscopy combined with fluorescent light microscopy (Briegel *et al.*, 2008). Although the study focused on only one MCP class, the method used in the study is promising for application to bacteria with multiple MCP classes, such as *Geobacter*. In the case of *G. sulfurreducens*, since there are 4 major MCP classes, 4 fusions with a fluorescent protein at the C-terminus of a representative of each class should be designed. The clusters of MCPs visualized by cryo-EM could be correlated with fluorescent cells to identify the cellular location of each MCP class.

2. Fluorescent protein techniques to study the segregation of MCP class.
    - Two MCPs of different classes (e.g 34H and 44H) could be fused with 2 different fluorescent proteins, and expressed individually and together in the *mcp*-negative background of *E. coli* (e.g. UU1250). Locations of the fusion proteins could be visualized under the microscope with specific filters for each fluorescent dye. We speculate that MCPs of different classes will segregate into different locations in

the cells. A potential problem that may be encountered is that MCPs from *Geobacter* may be unable to form a complex with CheA and CheW of *E. coli*, and the fluorescent signals therefore may not be detected, or may be spreading all over the cell. In that case, a screening method may be needed, using the template system (described below). The principle of the method is that if the signal complex is formed (between MCP, CheA, and CheW), there will be a marked difference in CheA activity compared to the activity of CheA itself (Shrout *et al.*, 2003). In parallel, this method could be used in *Geobacter* cells, in which the genes of two MCPs from different classes are replaced by genes with fluorescent fusions at the C-termini.

3. Template system to screen for MCPs' CheA and CheW partners.
    - Previous studies demonstrated that using a template system that contains 'artificial membrane' incubated with purified cytoplasmic fragments of MCPs, CheA and CheW at certain ratios could be used to study chemotaxis signaling complexes (Shrout et al., 2003).
4. *In vitro* methylation assay of mixed MCPs to test whether different classes of MCP could form a complex.
    - In *E. coli* it has been well understood that all five MCPs (of the same class) form a complex. Only two MCPs (Tsr and Tar) have the pentapeptide motif that is essential for methyltransferase CheR binding and therefore the methylation of these receptors. The other three MCPs do not have the pentapeptide motif and their methylation occurs due to the so-called inter-dimer methylation process (Li et al., 1997). The binding of CheR to MCPs containing a pentapeptide motif could

also help in transferring methyl groups to an MCP without a pentapeptide motif. We predict that the inter-dimer methylation process is class-restricted: i.e., only MCPs in the same class can be methylated by indirect association with CheR. The process, therefore, could be useful as an indirect way to test whether MCPs of the same class form cluster.

Examination of whether a CheR group is MCP class-specific and MCP tether-specific

The semi-conserved peptide at the C-terminus of some MCPs plays an important role in mediating efficient receptor methylation by enhancing the contact between substrate (MCP) and enzyme (CheR) via the β?-subdomain of this enzyme. We identified three groups of CheR in *Geobacter* based on: (1) the evaluation of conserved residues in the β?-subdomain; (2) the *che* cluster in which genes encoding CheR and MCPs are found; and (3) characteristics of the tether domain of MCPs encoded in the same *che* cluster with CheR. We postulated that each group of CheR is specific to a class of MCP with its typical tether domain. It was the first time this observation was reported, and experimental data are needed to support or rule out the hypothesis (Chapter 2). The following experiments could help to clarify the observation. They are designed to test whether (1) CheRs of group A specifically transfer methyl groups to class 36H MCPs with a xWxxF tether, but CheRs of group B or C do not; (b) CheRs of group B specifically transfer methyl groups to class 34H MCPs with a xFxxF tether, but CheRs of group C or A do not. The *in vitro* methylation assays will be carried out which use concentrated purified MCP and CheR, following the strategy previously described (Chervitz *et al.*, 1995):

- MCP of class 36H (with consensus xWxxF) and CheR in group A

- MCP of class of 34H (with consensus xFxxF) and CheR in group B
- MCP of class of 36H with mutation of the pentapeptide to make the tether like 34H, and CheR in group A, and then CheR in group B
- MCP class of 34H with mutation of the pentapeptide to make the tether like 36H, and CheR in group B, and then CheR in group A.
- MCP of class 34H or class 36H and CheR in group C

Completion of the understanding of gene regulation by the Che5 pathway

In Chapter 3, we showed evidences that the Che5 pathway, together with one class of MCP (class 40+24H) regulates the expression of OmcS, OmcZ and other genes involved in making extracellular materials. The model for the Che5 pathway was proposed to explain our data. However, there are questions that need to be answered in order to understand better the pathway depicted in Figure 36. These questions are: how does CheA5 activity regulate gene expression? And what are the signals to which the Che5 pathway responds?

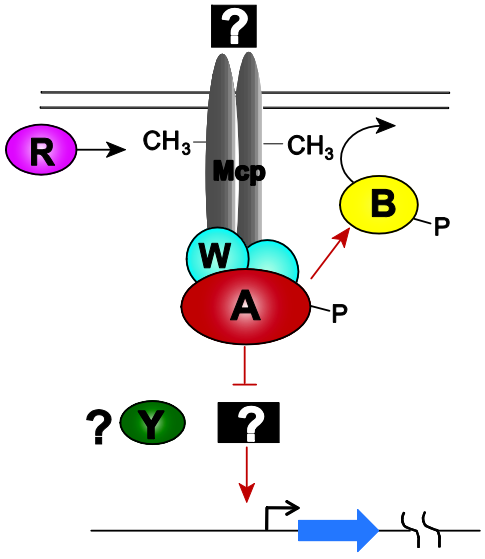


Figure 36. Che5 pathway model with questions to be addressed

Towards understanding how CheA5 activity regulate gene expression

The deletion of *cheA5*, a gene encoding an autophosphorylating histidine kinase, leads to the upregulation of *omcS* and other genes, suggesting that CheA5 activity inhibits expression of those genes. However, how CheA5 activity regulates gene expression is unclear, and its cognate response regulator is not yet identified. In *E. coli*, CheA transfers phosphate to CheY, and it is phosphorylated CheY (CheY~P) that interacts with FliM in the motor to regulate flagellar rotation and change the pattern of cell movement. Deletion of either *cheA* or *cheY* in *E. coli* makes the cell smooth swimming-biased (flagella rotate counter clock wise). There are three *cheY* genes in the *che5* cluster, but a single deletion mutant and double mutants downregulate OmcS, the opposite phenotype of a Δ*cheA* mutant (Chapter 3). This suggests the possibility that, in the Che5 pathway, CheA transfers phosphate to CheYs, but unphosphorylated CheY (most likely CheY5c) actually interacts with the target rather than CheY~P as seen in *E. coli*. This hypothesis could be tested by: (1) an *in vitro* phosphorylation assay to examine which CheY is phosphorylated by CheA; and (2) overexpression of CheY5c in *G. sulfurreducens*. It is predicted that the imbalance (bias toward the presence of more CheY) will increase OmcS.

Identification of signaling molecules of the Che5 pathway

The Che5 pathway regulates the expression of more than 170 genes, and *omcS* is mostly affected in the *che* mutants (Chapter 3). When CheA is inactivated, OmcS is upregulated, and when CheA is activated OmcS is downregulated, suggesting that in response to an increased concentration of attractants the cells will increase OmcS expression, and on the other hand, when the cells detect an increase in repellents they will

produce less OmcS. The amount of OmcS expression could be relatively quantified using a reporter gene such as the green fluorescent protein gene fused so that it is under regulation of the *omcS* promoter. This method could be used in a 96 well plate with a 96 well plate reader to screen a large number of signal molecules in a high-throughput mode. If a molecule is an attractant, the fluorescent signal will increase compared to the WT; on the other hand, if a molecule is a repellent, the signal will be lower than the WT. For longer readout with stable results, strains that are not able to adapt should be used: a Δ*cheB5* strain may be used to identify attractants, and a Δ*cheR5* strain may be used to identify repellents. The ideas could be simplified in the following Figure 37.

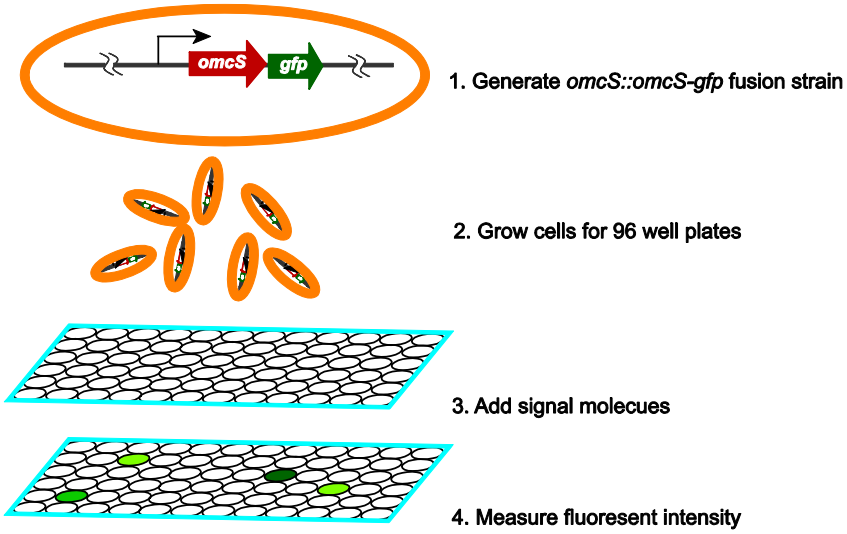


Figure 37. Strategy to identify signal molecules of the Che5 pathway

Further investigation of the signaling pathway of the *che1* cluster of *G. sulfurreducens* KN400

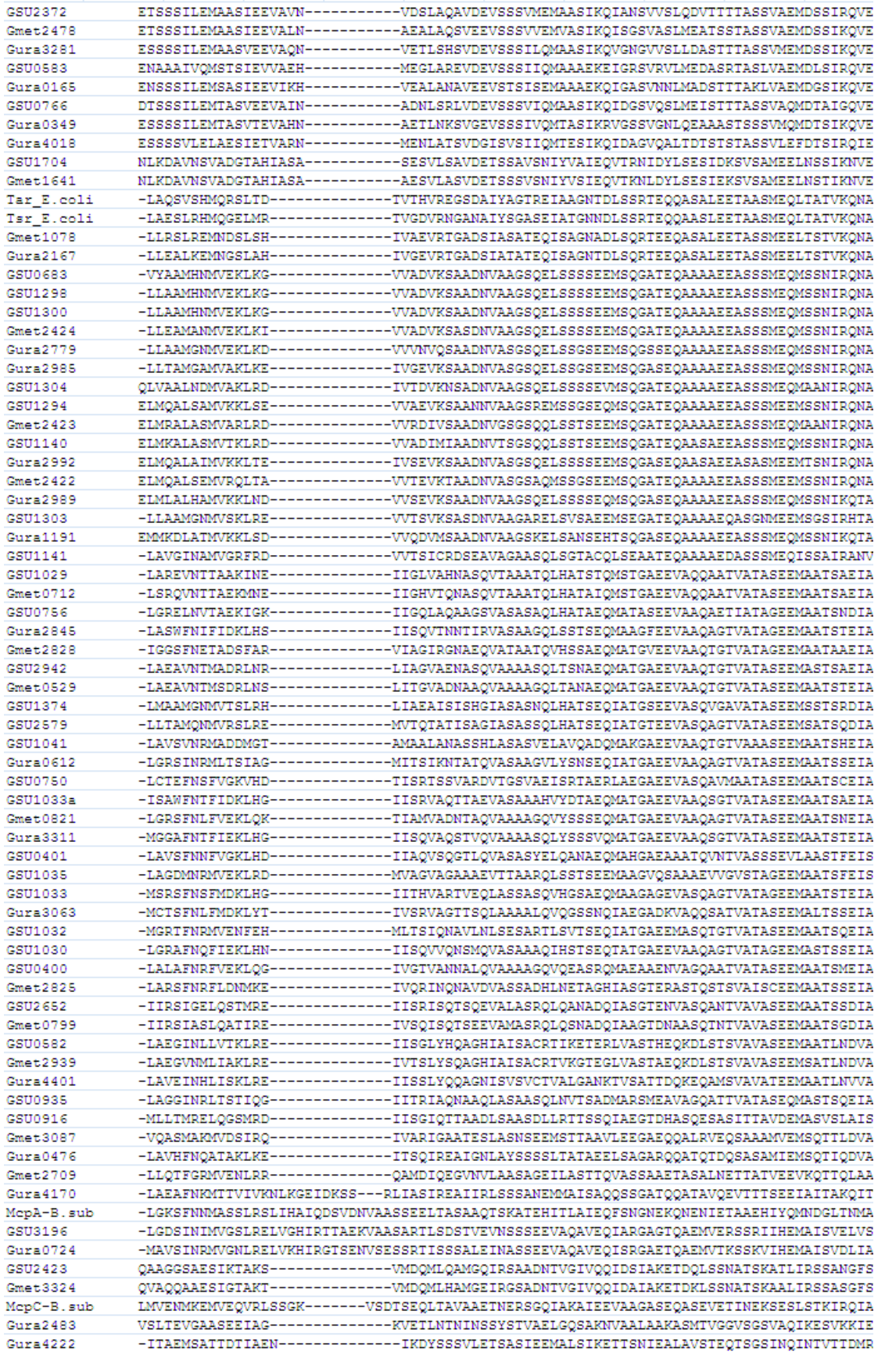
Our preliminary data (Chapter 4) showed that deletion of *gsu0296*, one of the four *cheA* genes of *G. sulfurreducens*, impaired its chemotaxis, suggesting that GSU0296 together with proteins encoded by the *che1* cluster regulates chemotaxis in *G. sulfurreducens* KN400. The following tasks are proposed in order to understand better the mechanism of the Che1 pathway:

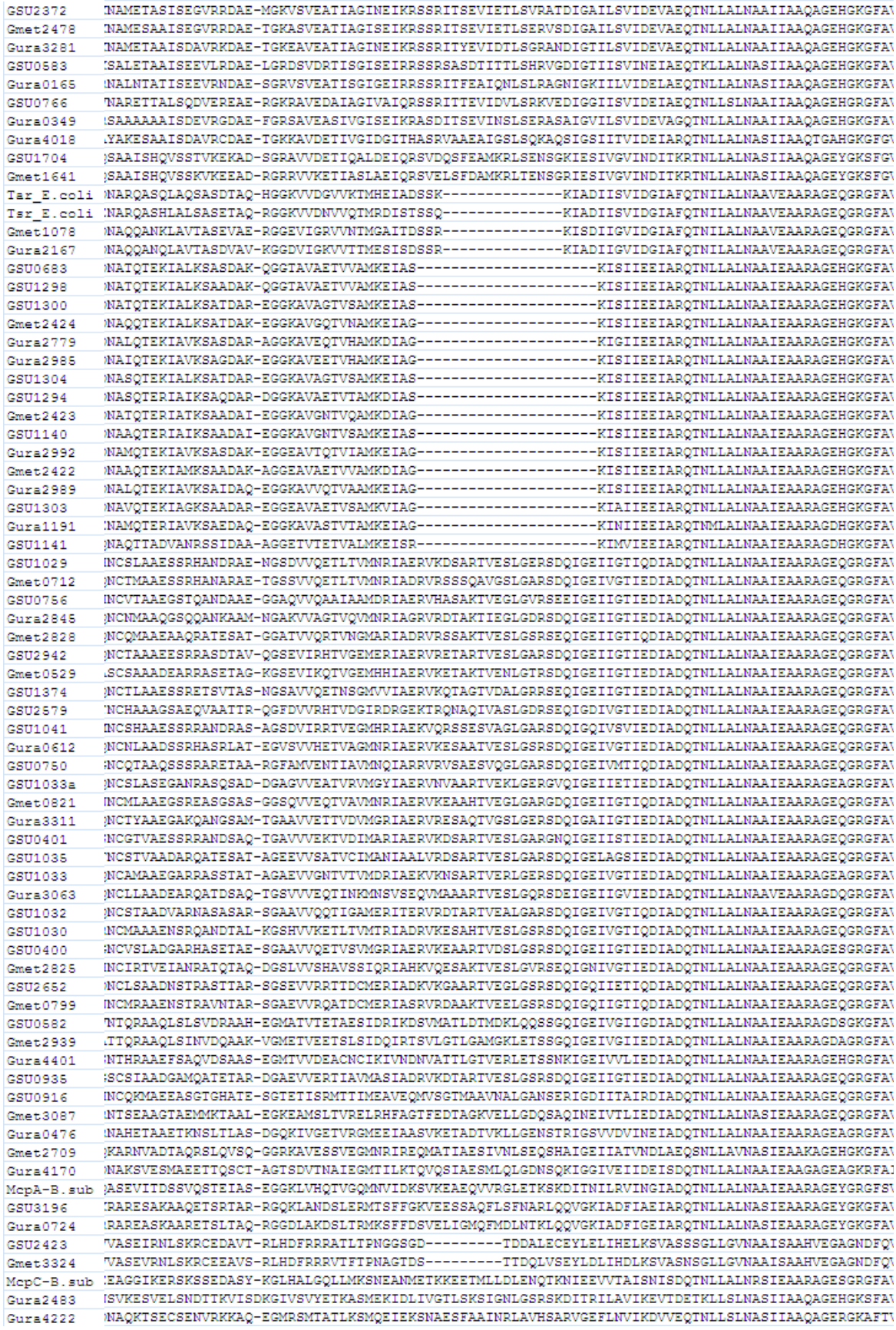
1. A complementation assay should be added to be sure that the phenotype observed is due to the absence of the gene, and not due to polar effects or unexpected secondary mutations.
2. Confirm that GSU0296 works with other proteins encoded by the *che1* cluster.
3. Identify the functions of each gene product in the *che1* cluster, in particular the function of the non-*che* gene, the HD domain protein.
4. Identify chemoattractants and/or repellents for the Che1 pathway.
5. Identify MCPs that signal through the Che1 pathway.

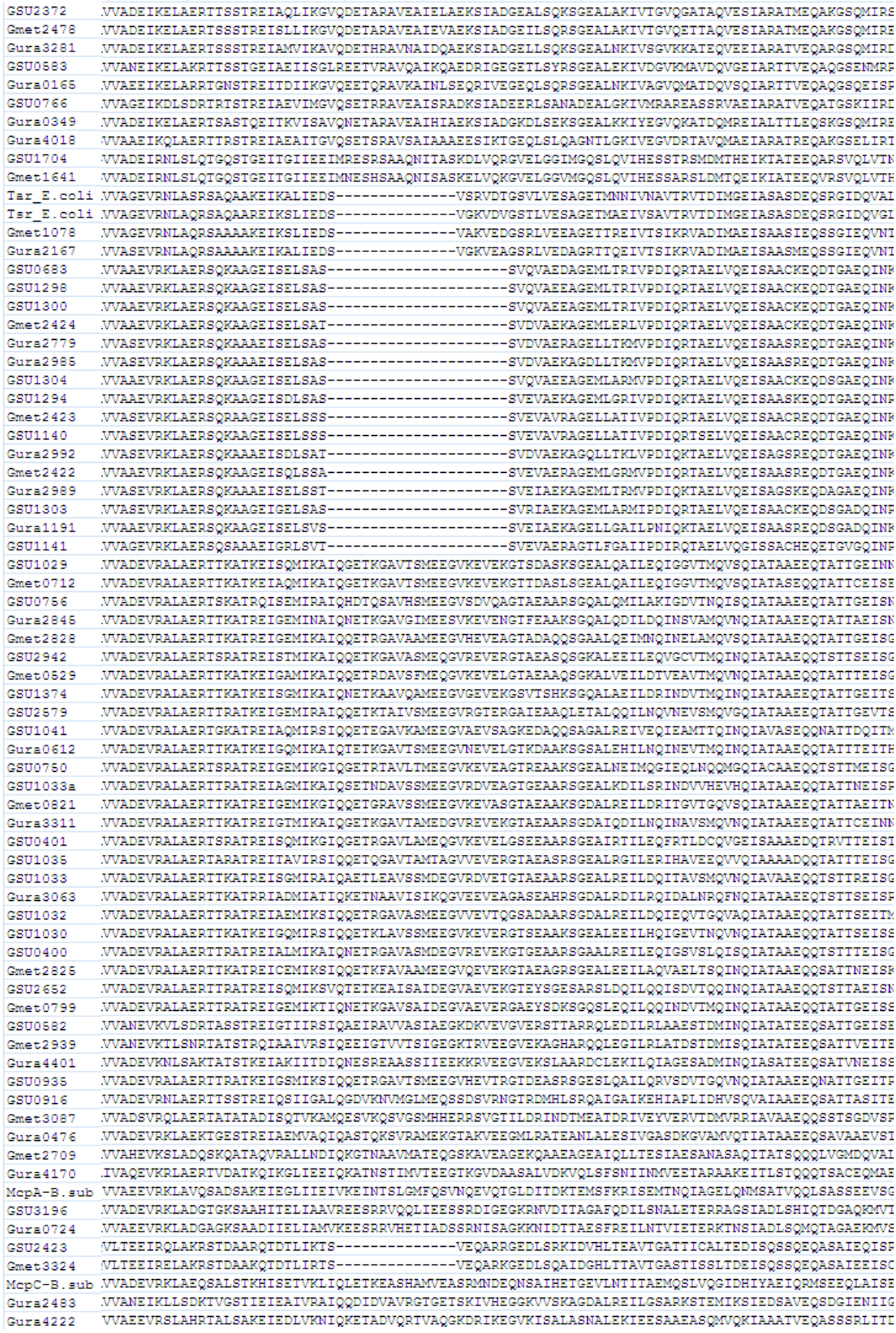
Test the impact of a chemotaxis defect on bioremediation (*in situ* or in the lab setup), as described in a previous study (Singh, 2008).

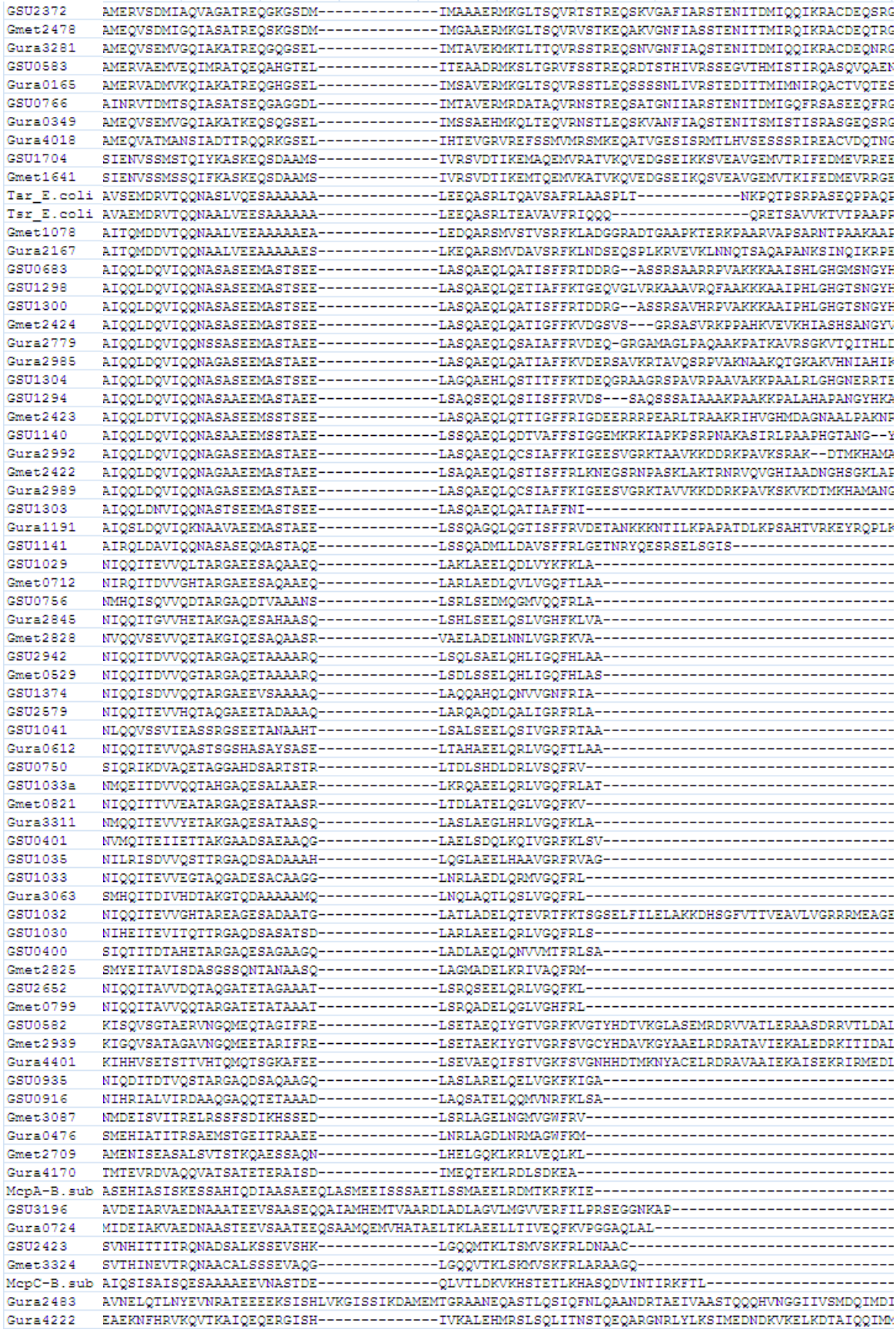
APPENDIX A

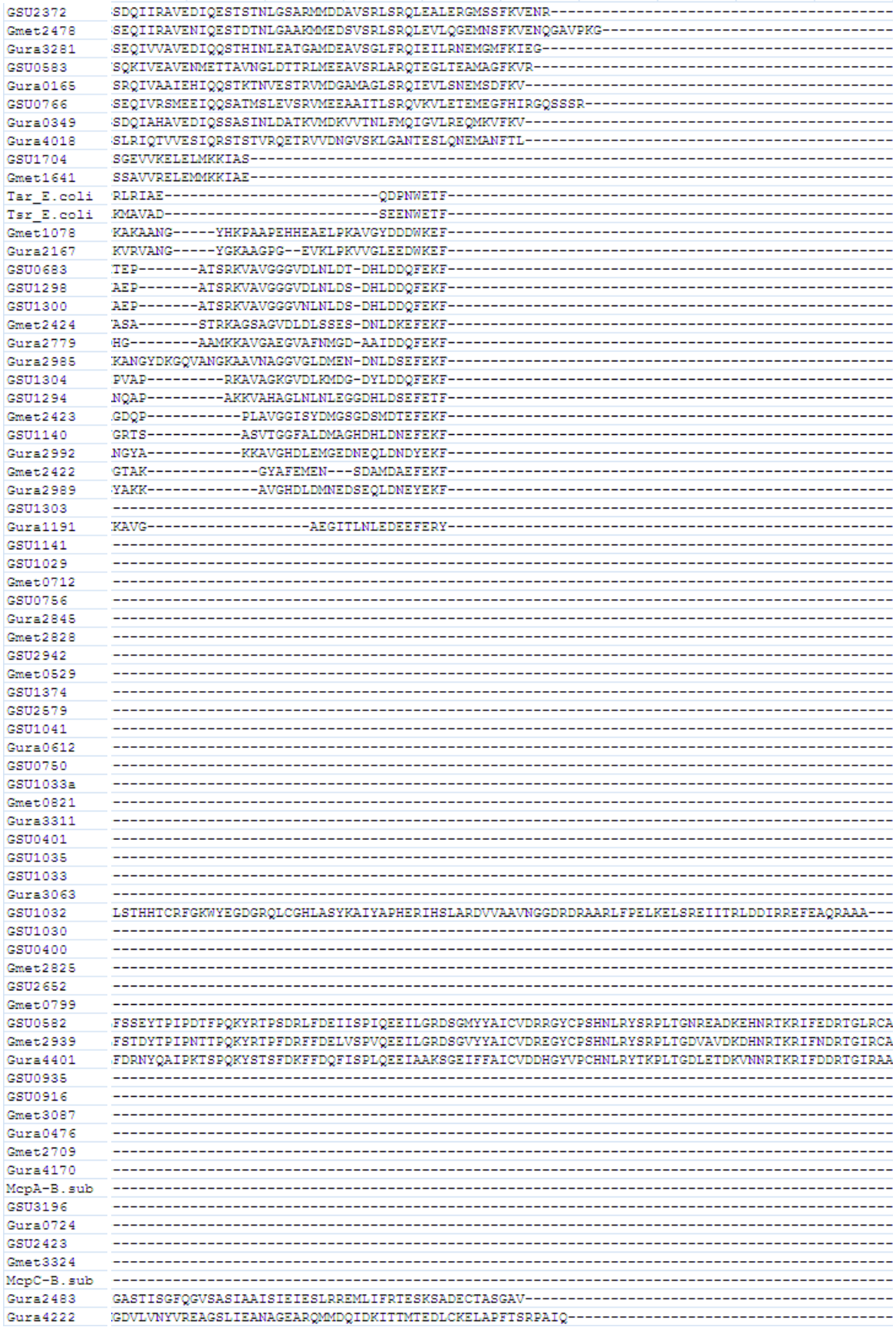
SEQUENCE ALIGNMENT OF CYTOPLASMIC DOMAINS FROM GEOBACTER SPECIES' MCPs


GSU2372 ETSSSILEMAASIEEVAVN----------VDSLAQAVDEVSSSVMEMAASIKQIANSVVSLQDVTTTTASSVAEMDSSIRQVE
Gmet2478 ETSSSILEMAASIEEVALN----------AEALAQSVEEVSSSVVEMVASIKQISGSVASLMEATSSTASSVAEMDSSIKQVE
Gura3281 ESSSSILEMAASVEEVAQN----------VETLSHSVDEVSSSILQMAASIKQVGNGVVSLLDASTTTASSVMEMDSSIKQVE
GSU0583 ENAAAIVQMSTSIEVVAEH----------MEGLAREVDEVSSSIIQMAAAEKEIGRSVRVLMEDASRTASLVAEMDLSIRQVE
Gura0165 ENSSSILEMSASIEEVIKH----------VEALANAVEEVSTSISEMAAAEKQIGASVNNLMADSTTTAKLVAEMDGSIKQVE
GSU0766 DTSSSILEMTASVEEVAIN----------ADNLSRLVDEVSSSVIQMAASIKQIDGSVQSLMEISTTTASSVAQMDIAIGQVE
Gura0349 ESSSSILEMTASVTEVAHN----------AETLNKSVGEVSSSIVQMTASIKRVGSSVGNLQEAAASTSSSVMQMDTSIKQVE
Gura4018 ESSSSVLELAESIETVARN----------MENLATSVDGISVSIIQMTESIKQIDAGVQALTDTSTSTASSVLEFDTSIRQIE
GSU1704 NLKDAVNSVADGTAHIASA----------SESVLSAVDETSSAVSNIYVAIEQVTRNIDYLSESIDKSVSAMEELNSSIKNVE
Gmet1641 NLKDAVNSVADGTAHIASA----------AESVLASVDETSSSVSNIYVSIEQVTKNLDYLSESIEKSVSAMEELNSTIKNVE
Tar_E.coli -LAQSVSHMQRSLTD----------TVTHVREGSDAIYAGTREIAAGNTDLSSRTEQQASALEETAASMEQLTATVKQNA
Tsr_E.coli -LAESLRHMQGELMR----------TVGDVRNGANAIYSGASEIATGNNDLSSRTEQQAASLEETAASMEQLTATVKQNA
Gmet1078 -LLRSLREMNDSLSH----------IVAEVRTGADSIASATEQISAGNADLSQRTEEQASALEETASSMEELTSTVKQNA
Gura2167 -LLEALKEMNGSLAH----------IVGEVRTGADSIATATEQISAGNTDLSQRTEEQASALEETASSMEELTSTVKQNA
GSU0683 -VYAAMHNMVEKLKG----------VVADVKSAADNVAAGSQELSSSSEEMSQGATEQAAAAEEASSSMEQMSSNIRQNA
GSU1298 -LLAAMHNMVEKLKG----------VVADVKSAADNVAAGSQELSSSSEEMSQGATEQAAAAEEASSSMEQMSSNIRQNA
GSU1300 -LLAAMHNMVEKLKG----------VVADVKSAADNVAAGSQELSSSSEEMSQGATEQAAAAEEASSSMEQMSSNIRQNA
Gmet2424 -LLEAMANMVEKLKI----------VVADVKSASDNVAAGSQELSSSSEEMSQGATEQAAAAEEASSSMEQMSSNIRQNA
Gura2779 -LLAAMGNMVEKLKD----------VVVNVQSAADNVASGSQELSSGSEEMSQGSSEQAAAAEEASSSMEQMSSNIRQNA
Gura2985 -LLTAMGAMVAKLKE----------IVGEVKSAADNVASGSQELSSGSEEMSQGASEQAAAAEEASSSMEQMSSNIRQNA
GSU1304 QLVAALNDMVAKLRD----------IVTDVKNSADNVAAGSQELSSSSEVMSQGATEQAAAAEEASSSMEQMAANIRQNA
GSU1294 ELMQALSAMVKKLSE----------VVAEVKSAANNVAAGSREMSSGSEQMSQGATEQAAAAEEASSSMEEMSSNIRQNA
Gmet2423 ELMRALASMVARLRD----------VVRDIVSAADNVGSGSQQLSSTSEEMSQGATEQAAAAEEASSSMEQMAANIRQNA
GSU1140 ELMKALASMVTKLRD----------VVADIMIAADNVTSGSQQLSSTSEEMSQGATEQAASAEEASSSMEQMSSNIRQNA
Gura2992 ELMQALAIMVKKLTE----------IVSEVKSAADNVASGSQELSSSSEEMSQGASEQAASAEEASASMEEMTSNIRQNA
Gmet2422 ELMQALSEMVRQLTA----------VVTEVKTAADNVASGSAQMSSGSEEMSQGATEQAAAAEEASSSMEEMSSNIRQNA
Gura2989 ELMLALHAMVKKLND----------VVSEVKSAADNVAAGSQELSSSSEQMSQGASEQAAAAEEASSSMEQMSSNIKQTA
GSU1303 -LLAAMGNMVSKLRE----------VVTSVKSASDNVAAGARELSVSAEEMSEGATEQAAAAEQASGNMEEMSGSIRHTA
Gura1191 EMMKDLATMVKKLSD----------VVQDVMSAADNVAAGSKELSANSEHTSQGASEQAAAAEEASSSMEQMSSNIKQTA
GSU1141 -LAVGINAMVGRFRD----------VVTSICRDSEAVAGAASQLSGTACQLSEAATEQAAAAEDASSSMEQISSAIRANV
GSU1029 -LAREVNTTAAKINE----------IIGLVAHNASQVTAAATQLHATSTQMSTGAEEVAQQAATVATASEEMAATSAEIA
Gmet0712 -LSRQVNTTAEKMNE----------IIGHVTQNASQVTAAATQLHATAIQMSTGAEEVAQQAATVATASEEMAATSAEIA
GSU0756 -LGRELNVTAEKIGK----------IIGQLAQAAGSVASASAQLHATAEQMATASEEVAAQAETIATAGEEMAATSNDIA
Gura2845 -LASWFNIFIDKLHS----------IISQVTNNTIRVASAAGQLSSTSEQMAAGFEEVAAQAGTVATAGEEMAATSTEIA
Gmet2828 -IGGSFNETADSFAR----------VIAGIRGNAEQVATAATQVHSSAEQMATGVEEVAAQTGTVATAGEEMAATAAEIA
GSU2942 -LAEAVNTMADRLNR----------LIAGVAENASQVAAAASQLTSNAEQMATGAEEVAAQTGTVATASEEMASTSAEIA
Gmet0529 -LAEAVNTMSDRLNS----------LITGVADNAAQVAAAAGQLTANAEQMATGAEEVAAQTGTVATASEEMAATSTEIA
GSU1374 -LMAAMGNMVTSLRH----------LIAEAISISHGIASASNQLHATSEQIATGSEEVASQVGAVATASEEMSSTSRDIA
GSU2579 -LLTAMQNMVRSLRE----------MVTQTATISAGIASASSQLHATSEQIATGTEEVASQAGTVATASEEMSATSQDIA
GSU1041 -LAVSVNRMADDMGT----------AMAALANASSHLASASVELAVQADQMAKGAEEVAAQTGTVAAASEEMAATSHEIA
Gura0612 -LGRSINRMLTSIAG----------MITSIKNTATQVASAAGVLYSNSEQIATGAEEVAAQAGTVATASEEMAATSSEIA
GSU0750 -LCTEFNSFVGKVHD----------TISRTSSVARDVTGSVAEISRTAERLAEGAEEVASQAVMAATASEEMAATSCEIA
GSU1033a -ISAWFNTFIDKLHG----------IISRVAQTTAEVASAAAHVYDTAEQMATGAEEVAAQSGTVATASEEMAATSAEIA
Gmet0821 -LGRSFNLFVEKLQK----------TIAMVADNTAQVAAAAGQVYSSSEQMATGAEEVAAQAGTVATASEEMAATSNEIA
Gura3311 -MGGAFNTFIEKLHG----------IISQVAQSTVQVAAAASQLYSSSVQMATGAEEVAAQSGTVATASEEMAATSTEIA
GSU0401 -LAVSFNNFVGKLHD----------IIAQVSQGTLQVASASYELQANAEQMAHGAEAAATQVNTVASSSEVLAASTFEIS
GSU1035 -LAGDMNRMVEKLRD----------MVAGVAGAAAEVTTAARQLSSTSEEMAAGVQSAAAEVVGVSTAGEEMAATSFEIS
GSU1033 -MSRSFNSFMDKLHG----------IITHVARTVEQLASSASQVHGSAEQMAAGAGEVASQAGTVATAGEEMAATSTEIA
Gura3063 -MCTSFNLFMDKLYT----------IVSRVAGTTSQLAAAALQVQGSSNQIAEGADKVAQQSATVATASEEMALTSSEIA
GSU1032 -MGRTFNRMVENFEH----------MLTSIQNAVLNLSESARTLSVTSEQIATGAEEMASQTGTVATASEEMAATSQEIA
GSU1030 -LGRAFNQFIEKLHN----------IISQVVQNSMQVASAAAQIHSTSEQTATGAEEVAAQAGTVATAGEEMASTSSEIA
GSU0400 -LALAFNRFVEKLQG----------IVGTVANNALQVAAAAGQVQEASRQMAEAAENVAGQAATVATASEEMAATSMEIA
Gmet2825 -LARSFNRFLDNMKE----------IVQRINQNAVDVASSADHLNETAGHIASGTERASTQSTSVAISCEEMAATSSEIA
GSU2652 -IIRSIGELQSTMRE----------IISRISQTSQEVALASRQLQANADQIASGTENVASQANTVAVASEEMAATSSDIA
Gmet0799 -IIRSIASLQATIRE----------IVSQISQTSEEVAMASRQLQSNADQIAAGTDNAASQTNTVAVASEEMAATSGDIA
GSU0582 -LAEGINLLVTKLRE----------IISGLYHQAGHIAISACRTIKETERLVASTHEQKDLSTSVAVASEEMAATLNDVA
Gmet2939 -LAEGVNMLIAKLRE----------IVTSLYSQAGHIAISACRTVKGTEGLVASTAEQKDLSTSVAVASEEMSATLNDVA
Gura4401 -LAVEINHLISKLRE----------IISSLYQQAGNISVSVCTVALGANKTVSATTDQKEQAMSVAVATEEMAATLNVVA
GSU0935 -LAGGINRLTSTIQG----------IITRIAQNAAQLASAASQLNVTSADMARSMEAVAGQATTVATASEQMASTSQEIA
GSU0916 -MLLTMRELQGSMRD----------IISGIQTTAADLSAASDLLRTTSSQIAEGTDHASQESASITTAVDEMASVSLAIS
Gmet3087 -VQASMAKMVDSIRQ----------IVARIGAATESLASNSEEMSTTAAVLEEGAEQQALRVEQSAAAMVEMSQTTLDVA
Gura0476 -LAVHFNQATAKLKE----------ITSQIREAIGNLAYSSSSLTATAEELSAGARQQATQTDQSASAMIEMSQTIQDVA
Gmet2709 -LLQTFGRMVENLRR----------QAMDIQEGVNVLAASAGEILASTTQVASSAAETASALNETTATVEEVKQTTQLAA
Gura4170 -LAEAFNKMTTVIVKNLKGEIDKSS---RLIASIREAIIRLSSSANEMMAISAQQSSGATQQATAVQEVTTTSEEIAITAKQIT
McpA-B.sub -LGKSFNNMASSLRSLIHAIQDSVDNVAASSEELTASAAQTSKATEHITLAIEQFSNGNEKQNENIETAAEHIYQMNDGLTNMA
GSU3196 -LGDSINIMVGSLRELVGHIRTTAEKVAASARTLSDSTVEVNSSSEEVAQAVEQIARGAGTQAEMVERSSRIIHEMAISVELVS
Gura0724 -MAVSINRMVGNLRELVKHIRGTSENVSESSRTISSSALEINASSEEVAQAVEQISRGAETQAEMVTKSSKVIHEMAISVDLIA
GSU2423 QAAGGSAESIKTAKS----------VMDQMLQAMGQIRSAADNTVGIVQQIDSIAKETDQLSSNATSKATLIRSSANGFS
Gmet3324 QVAQQAAESIGTAKT----------VMDQMLHAMGEIRGSADNTVGIVQQIDAIAKETDKLSSNATSKAALIRSSASGFS
McpC-B.sub LMVENMKEMVEQVRLSSGK-------VSDTSEQLTAVAAETNERSGQIAKAIEEVAAGASEQASEVETINEKSESLSTKIRQIA
Gura2483 VSLTEVGAASEEIAG----------KVETLNTNINSSYSTVAELGQSAKNVAALAAKASMTVGGVSGSVAQIKESVKKIE
Gura4222 -ITAEMSATTDTIAEN----------IKDYSSSVLETSASIEEMALSIKETTSNIEALAVSTEQTSGSINQINTVTTDMR


GSU2372
Gmet2478
Gura3281
GSU0583
Gura0165
GSU0766
Gura0349
Gura4018
GSU1704
Gmet1641
Tar_E.coli
Tsr_E.coli
Gmet1078
Gura2167
GSU0683
GSU1298
GSU1300
Gmet2424
Gura2779
Gura2985
GSU1304
GSU1294
Gmet2423
GSU1140
Gura2992
Gmet2422
Gura2989
GSU1303
Gura1191
GSU1141
GSU1029
Gmet0712
GSU0756
Gura2845
Gmet2828
GSU2942
Gmet0529
GSU1374
GSU2579
GSU1041
Gura0612
GSU0750
GSU1033a
Gmet0821
Gura3311
GSU0401
GSU1035
GSU1033
Gura3063
GSU1032
GSU1030
GSU0400
Gmet2825
GSU2652
Gmet0799
GSU0582
Gmet2939
Gura4401
GSU0935
GSU0916
Gmet3087
Gura0476
Gmet2709
Gura4170
McpA-B.sub
GSU3196
Gura0724
GSU2423
Gmet3324
McpC-B.sub
Gura2483
Gura4222




GSU2372 AMERVSDMIAQVAGATREQGKGSDM----------IMAAAERMKGLTSQVRTSTREQSKVGAFIARSTENITDMIQQIKRACDEQSRG
Gmet2478 AMEQVSDMIGQIASATREQSKGSDM----------IMGAAERMKGLTSQVRVSTKEQAKVGNFIASSTENITTMIRQIKRACDEQTRG
Gura3281 AMEQVSEMVGQIAKATREQGQGSEL----------IMTAVEKMKTLTTQVRSSTREQSNVGNFIAQSTENITDMIQQIKRACDEQNRG
GSU0583 AMERVAEMVEQIMRATQEQAHGTEL----------ITEAADRMKSLTGRVFSSTREQRDTSTHIVRSSEGVTHMISTIRQASQVQAEN
Gura0165 AMERVADMVKQIAKATREQGHGSEL----------IMSAVERMKGLTSQVRSSTLEQSSSSNLIVRSTEDITTMIMNIRQACTVQTES
GSU0766 AINRVTDMTSQIASATSEQGAGGDL----------IMTAVERMRDATAQVRNSTREQSATGNIIARSTENITDMIGQFRSASEEQFRG
Gura0349 AMEQVSEMVGQIAKATKEQSQGSEL----------IMSSAEHMKQLTEQVRNSTLEQSKVANFIAQSTENITSMISTISRASGEQSRG
Gura4018 AMEQVATMANSIADTTRQQRKGSEL----------IHTEVGRVREFSSMVMRSMKEQATVGESISRMTLHVSESSSRIREACVDQTNG
GSU1704 SIENVSSMSTQIYKASKEQSDAAMS----------IVRSVDTIKEMAQEMVRATVKQVEDGSEIKKSVEAVGEMVTRIFEDMEVRREE
Gmet1641 SIENVSSMSSQIFKASKEQSDAAMS----------IVRSVDTIKEMTQEMVKATVKQVEDGSEIKQSVEAVGEMVTKIFEDMEVRRGE
Tar_E.coli AVSEMDRVTQQNASLVQESAAAAAA----------LEEQASRLTQAVSAFRLAASPLT----------NKPQTPSRPASEQPPAQP
Tsr_E.coli AVAEMDRVTQQNAALVEESAAAAAA----------LEEQASRLTEAVAVFRIQQQ----------QRETSAVVKTVTPAAPP
Gmet1078 AITQMDDVTQQNAALVEEAAAAAEA----------LEDQARSMVSTVSRFKLADGGRADTGAAPKTERKPAARVAPSARNTPAAKAAP
Gura2167 AITQMDDVTQQNAALVEEAAAAAES----------LKEQARSMVDAVSRFKLNDSEQSPLKRVEVKLNNQTSAQAPANKSINQIKRPE
GSU0683 AIQQLDQVIQQNASASEEMASTSEE----------LASQAEQLQATISFFRTDDRG--ASSRSAARRPVAKKKAAISHLGHGMSNGYH
GSU1298 AIQQLDQVIQQNASASEEMASTSEE----------LASQAEQLQETIAFFKTGEQVGLVRKAAAVRQFAAKKKAAIPHLGHGTSNGYH
GSU1300 AIQQLDQVIQQNASASEEMASTSEE----------LASQAEQLQATISFFRTDDRG--ASSRSAVHRPVAKKKAAIPHLGHGTSNGYH
Gmet2424 AIQQLDQVIQQNASASEEMASTSEE----------LASQAEQLQATIGFFKVDGSVS---GRSASVRKPPAHKVEVKHIASHSANGYV
Gura2779 AIQQLDQVIQQNSSASEEMASTAEE----------LASQAEQLQSAIAFFRVDEQ-GRGAMAGLPAQAAKPATKAVRSGKVTQITHLD
Gura2985 AIQQLDQVIQQNAGASEEMASTAEE----------LASQAEQLQATIAFFKVDERSAVKRTAVQSRPVAKNAAKQTGKAKVHNIAHIR
GSU1304 AIQQLDQVIQQNASASEEMASTSEE----------LAGQAEHLQSTITFFKTDEQGRAAGRSPAVRPAAVAKKPAALRLGHGNERRTE
GSU1294 AIQQLDQVIQQNASAAEEMASTAEE----------LSAQSEQLQSIISFFRVDS---SAQSSSAIAAAKPAAKKPALAHAPANGYHKA
Gmet2423 AIQQLDTVIQQNASASEEMSSTSEE----------LASQAEQLQTTIGFFRIGDEERRRPEARLTRAAKRIHVGHMDAGNAALPAKNP
GSU1140 AIQQLDQVIQQNASAAEEMSSTAEE----------LSSQAEQLQDTVAFFSIGGEMKRKIAPKPSRPNAKASIRLPAAPHGTANG--Y
Gura2992 AIQQLDQVIQQNAGASEEMASTAEE----------LASQAEQLQCSIAFFKIGEESVGRKTAAVKKDDRKPAVKSRAK--DTMKHAMA
Gmet2422 AIQQLDQVIQQNAGAAEEMASTAEE----------LSAQAEQLQSTISFFRLKNEGSRNPASKLAKTRNRVQVGHIAADNGHSGKLAE
Gura2989 AIQQLDQVIQQNAGASEEMASTAEE----------LASQAEQLQCSIAFFKIGEESVGRKTAVVKKDDRKPAVKSKVKDTMKHAMANG
GSU1303 AIQQLDNVIQQNASTSEEMASTSEE----------LASQAEQLQATIAFFNI----------
Gura1191 AIQSLDQVIQKNAAVAEEMASTAEE----------LSSQAGQLQGTISFFRVDETANKKKNTILKPAPATDLKPSAHTVRKEYRQPLR
GSU1141 AIRQLDAVIQQNASASEQMASTAQE----------LSSQADMLLDAVSFFRLGETNRYQESRSELSGIS----------
GSU1029 NIQQITEVVQLTARGAEESAQAAEQ----------LAKLAEELQDLVYKFKLA----------
Gmet0712 NIRQITDVVGHTARGAEESAQAAEQ----------LARLAEDLQVLVGQFTLAA----------
GSU0756 NMHQISQVVQDTARGAQDTVAAANS----------LSRLSEDMQGMVQQFRLA----------
Gura2845 NIQQITGVVHETAKGAQESAHAASQ----------LSHLSEELQSLVGHFKLVA----------
Gmet2828 NVQQVSEVVQETAKGIQESAQAASR----------VAELADELNNLVGRFKVA----------
GSU2942 NIQQITDVVQQTARGAQETAAAARQ----------LSQLSAELQHLIGQFHLAA----------
Gmet0529 NIQQITDVVQGTARGAQETAAAARQ----------LSDLSSELQHLIGQFHLAS----------
GSU1374 NIQQISDVVQQTARGAEEVSAAAAQ----------LAQQAHQLQNVVGNFRIA----------
GSU2579 NIQQITEVVHQTAQGAEETADAAAQ----------LARQAQDLQALIGRFRLA----------
GSU1041 NLQQVSSVIEASSRGSEETANAAHT----------LSALSEELQSIVGRFRTAA----------
Gura0612 NIQQITEVVQASTSGSHASAYSASE----------LTAHAEELQRLVGQFTLAA----------
GSU0750 SIQRIKDVAQETAGGAHDSARTSTR----------LTDLSHDLDRLVSQFRV----------
GSU1033a NMQEITDVVQQTAHGAQESALAAER----------LKRQAEELQRLVGQFRLAT----------
Gmet0821 NIQQITTVVEATARGAQESATAASR----------LTDLATELQGLVGQFKV----------
Gura3311 NMQQITEVVYETAKGAQESATAASQ----------LASLAEGLHRLVGQFKLA----------
GSU0401 NVMQITEIIETTAKGAADSAEAAQG----------LAELSDQLKQIVGRFKLSV----------
GSU1035 NILRISDVVQSTTRGAQDSADAAAH----------LQGLAEELHAAVGRFRVAG----------
GSU1033 NIQQITEVVEGTAQGADESACAAGG----------LNRLAEDLQRMVGQFRL----------
Gura3063 SMHQITDIVHDTAKGTQDAAAAAMQ----------LNQLAQTLQSLVGQFRL----------
GSU1032 NIQQITEVVGHTAREAGESADAATG----------LATLADELQTEVRTFKTSGSELFILELAKKDHSGFVTTVEAVLVGRRRMEAGE
GSU1030 NIHEITEVITQTTRGAQDSASATSD----------LARLAEELQRLVGQFRLS----------
GSU0400 SIQTITDTAHETARGAQESAGAAGQ----------LADLAEQLQNVVMTFRLSA----------
Gmet2825 SMYEITAVISDASGSSQNTANAASQ----------LAGMADELKRIVAQFRM----------
GSU2652 NIQQITAVVDQTAQGATETAGAAAT----------LSRQSEELQRLVGQFKL----------
Gmet0799 NIQQITAVVQQTARGATETATAAAT----------LSRQADELQGLVGHFRL----------
GSU0582 KISQVSGTAERVNGQMEQTAGIFRE----------LSETAEQIYGTVGRFKVGTYHDTVKGLASEMRDRVVATLERAASDRRVTLDAI
Gmet2939 KIGQVSATAGAVNGQMEETARIFRE----------LSETAEKIYGTVGRFSVGCYHDAVKGYAAELRDRATAVIEKALEDRKITIDAI
Gura4401 KIHHVSETSTTVHTQMQTSGKAFEE----------LSEVAEQIFSTVGKFSVGNHHDTMKNYACELRDRAVAAIEKAISEKRIRMEDI
GSU0935 NIQDITDTVQSTARGAQDSAQAAGQ----------LASLARELQELVGKFKIGA----------
GSU0916 NIHRIALVIRDAAQGAQQTETAAAD----------LAQSATELQQMVNRFKLSA----------
Gmet3087 NMDEISVITRELRSSFSDIKHSSED----------LSRLAGELNGMVGWFRV----------
Gura0476 SMEHIATITRSAEMSTGEITRAAEE----------LNRLAGDLNRMAGWFKM----------
Gmet2709 AMENISEASALSVTSTKQAESSAQN----------LHELGQKLKRLVEQLKL----------
Gura4170 TMTEVRDVAQQVATSATETERAISD----------IMEQTEKLRDLSDKEA----------
McpA-B.sub ASEHIASISKESSAHIQDIAASAEEQLASMEEISSSAETLSSMAEELRDMTKRFKIE----------
GSU3196 AVDEIARVAEDNAAATEEVSAASEQQAIAMHEMTVAARDLADLAGVLMGVVERFILPRSEGGNKAP----------
Gura0724 MIDEIAKVAEDNAASTEEVSAATEEQSAAMQEMVHATAELTKLAEELLTIVEQFKVPGGAQLAL----------
GSU2423 SVNHITTITRQNADSALKSSEVSHK----------LGQQMTKLTSMVSKFRLDNAAC----------
Gmet3324 SVTHINEVTRQNAACALSSSEVAQG----------LGQQVTKLSKMVSKFRLARAAGQ----------
McpC-B.sub AIQSISAISQESAAAAEEVNASTDE----------QLVTLDKVKHSTETLKHASQDVINTIRKFTL----------
Gura2483 AVNELQTLNYEVNPATEEEEKSISHLVKGISSIKDAMEMTGRAANEQASTLQSIQFNLQAANDRTAEIVAASTQQQHVNGGIIVSMDQIMDI
Gura4222 EAEKNFHRVKQVTKAIQEQERGISH----------IVKALEHMRSLSQLITNSTQEQARGNRLYLKSIMEDNDKVKELKDTAIQQIMN


GSU2372 SDQIIRAVEDIQESTSTNLGSARMMDDAVSRLSRQLEALERGMSSFKVENR----------
Gmet2478 SEQIIRAVENIQESTDTNLGAAKMMEDSVSRLSRQLEVLQGEMNSFKVENQGAVPKG----------
Gura3281 SEQIVVAVEDIQQSTHINLEATGAMDEAVSGLFRQIEILRNEMGMFKIEG----------
GSU0583 SQKIVEAVENMETTAVNGLDTTRLMEEAVSRLARQTEGLTEAMAGFKVR----------
Gura0165 SRQIVAAIEHIQQSTKTNVESTRVMDGAMAGLSRQIEVLSNEMSDFKV----------
GSU0766 SEQIVRSMEEIQQSATMSLEVSRVMEEAAITLSRQVKVLETEMEGFHIRGQSSSR----------
Gura0349 SDQIAHAVEDIQSSASINLDATKVMDKVVTNLFMQIGVLREQMKVFKV----------
Gura4018 SLRIQTVVESIQRSTSTVRQETRVVDNGVSKLGANTESLQNEMANFTL----------
GSU1704 SGEVVKELELMKKIAS----------
Gmet1641 SSAVVRELEMMKKIAE----------
Tar_E.coli RLRIAE----------QDPNWETF----------
Tsr_E.coli KMAVAD----------SEENWETF----------
Gmet1078 KAKAANG-----YHKPAAPEHHEAELPKAVGYDDDWKEF----------
Gura2167 KVRVANG-----YGKAAGPG--EVKLPKVVGLEEDWKEF----------
GSU0683 TEP-------ATSRKVAVGGGVDLNLDT-DHLDDQFEKF----------
GSU1298 AEP-------ATSRKVAVGGGVDLNLDS-DHLDDQFEKF----------
GSU1300 AEP-------ATSRKVAVGGGVNLNLDS-DHLDDQFEKF----------
Gmet2424 ASA-------STRKAGSAGVDLDLSSES-DNLDKEFEKF----------
Gura2779 HG--------AAMKKAVGAEGVAFNMGD-AAIDDQFEKF----------
Gura2985 KANGYDKGQVANGKAAVNAGGVGLDMEN-DNLDSEFEKF----------
GSU1304 PVAP---------RKAVAGKGVDLKMDG-DYLDDQFEKF----------
GSU1294 NQAP---------AKKVAHAGLNLNLEGGDHLDSEFETF----------
Gmet2423 GDQP-----------PLAVGGISYDMGSGDSMDTEFEKF----------
GSU1140 GRTS-----------ASVTGGFALDMAGHDHLDNEFEKF----------
Gura2992 NGYA-----------KKAVGHDLEMGEDNEQLDNDYEKF----------
Gmet2422 GTAK-------------GYAFEMEN---SDAMDAEFEKF----------
Gura2989 YAKK-------------AVGHDLDMNEDSEQLDNEYEKF----------
GSU1303 ----------
Gura1191 KAVG-----------------AEGITLNLEDEEFERY----------
GSU1141 ----------
GSU1029 ----------
Gmet0712 ----------
GSU0756 ----------
Gura2845 ----------
Gmet2828 ----------
GSU2942 ----------
Gmet0529 ----------
GSU1374 ----------
GSU2579 ----------
GSU1041 ----------
Gura0612 ----------
GSU0750 ----------
GSU1033a ----------
Gmet0821 ----------
Gura3311 ----------
GSU0401 ----------
GSU1035 ----------
GSU1033 ----------
Gura3063 ----------
GSU1032 LSTHHTCRFGKWYEGDGRQLCGHLASYKAIYAPHERIHSLARDVVAAVNGGDRDRAARLFPELKELSREIITRLDDIRREFEAQRAAA---
GSU1030 ----------
GSU0400 ----------
Gmet2825 ----------
GSU2652 ----------
Gmet0799 ----------
GSU0582 FSSEYTPIPDTFPQKYRTPSDRLFDEIISPIQEEILGRDSGMYYAICVDRRGYCPSHNLRYSRPLTGNREADKEHNRTKRIFEDRTGLRCA
Gmet2939 FSTDYTPIPNTTPQKYRTPFDRFFDELVSPVQEEILGRDSGVYYAICVDREGYCPSHNLRYSRPLTGDVAVDKDHNRTKRIFNDRTGIRCA
Gura4401 FDRNYQAIPKTSPQKYSTSFDKFFDQFISPLQEEIAAKSGEIFFAICVDDHGYVPCHNLRYTKPLTGDLETDKVNNRTKRIFDDRTGIRAA
GSU0935 ----------
GSU0916 ----------
Gmet3087 ----------
Gura0476 ----------
Gmet2709 ----------
Gura4170 ----------
McpA-B.sub ----------
GSU3196 ----------
Gura0724 ----------
GSU2423 ----------
Gmet3324 ----------
McpC-B.sub ----------
Gura2483 GASTISGFQGVSASIAAISIEIESLRREMLIFRTESKSADECTASGAV----------
Gura4222 GDVLVNYVREAGSLIEANAGEARQMMDQIDKITTMTEDLCKELAPFTSRPAIQ----------

APPENDIX B

LIST OF PRIMERS USED IN CHAPTER 3

Purpose	Primers and sequences	Descriptions
Primers for creating mutation		
gsu0296::kan	HT55F 5'-GCTGGTGAGTTTCAAGCTAGAGGAAGAGG-3' HT55R 5'-CGATAAGCTTAATCGCCATGTGCTGCTCC-3'	To PCR *gsu0296* upstream
	HT42F 5'-GCTGCAAGCTTTGGGATGAATGTCAGCTAC-3' HT22R 5'-GCACTCTAGAGCGGTGGAATCGAAATCTCG-3'	To PCR Kan
	HT56F 5'-GCTATCTAGACCTGAGCGGACAACGAGC-3' HT56R 5'-GCTGATGTCCGTACCGACGATTTCGATG-3'	To PCR *gsu0296* downstream
gsu1290::kan	HT57F 5'-CGTCGATATTACCGGCCTGCAACTGC-3' HT57R 5'-GCATAAGCTTGTCCATTAGCCGAGCACC-3'	To PCR *gsu1290* upstream
	HT42F and HT22R as above	To PCR Kan
	HT58F 5'-GCAGTCTAGACCGGTAGAATGTGCTCTGC-3' HT58R 5'-GGTAGATCAGGTTGCATCTGGGGAGAACC-3'	To PCR *gsu1290* downstream
gsu2222::kan	HT61F 5'-CCATCCTCTACGGCGACAATCTGC-3' HT61R 5'-CGTAGTCGACGTGTTGGTCATGGAACC-3'	To PCR *gsu2222* upstream
	HT22F 5'-GCTAGTCGACTGGGATGAATGTCAGCTAC-3' HT22R as above	To PCR Kan
	HT62F' 5'-GCTATCTAGAGGCAGCAGGACATCGTCATC-3' HT62R' 5'-GCATCAGGTAGAGCGTTTCCGTGAG-3'	To PCR *gsu2222* downstream
gsu3199::kan	HT59F 5'-CGAGTGAACATCCGCGTTTCGAGGGATAC-3' HT59R 5'-GCTCAAGCTTGGACATGTCCATGTCTAGC-3'	To PCR *3199* upstream
	HT42F and HT22R as above	To PCR Kan
	HT60F' 5'-GCAGCTCTAGAGCAGAGGTATTCACCAAGC-3' HT60R' 5'-CGAGATCTCTAGCTCTGTGAGGAGGGACC-3'	To PCR *gsu3199* downstream
gsu2210::kan	HT149F 5'-GGTCATGGACGGTCTCAAGCTGGTGAG-3' HT149R 5'-GCTACAAGCTTCATGGCCTTCATGGTGACC-3'	To PCR *gsu2210* upstream
	HT42F and HT22R as above	To PCR Kan
	HT150F 5'-GCTAATCTAGAGCTGCGTTTCCCGTCACC-3' HT150R 5'-CGACGGATCTTCGGTGACCGTGAAG-3'	To PCR *gsu2210* downstream
gsu2212::kan	HT113F 5'-CGACCCTCGATTTTAGCGAAGTGCTCC-3' HT113R 5'-GCTAGAGCTCCATGATGAACGTCTCCTCAG-3'	To PCR *gsu2212* upstream
	HT114F 5'-GCTCGAGCTCTGGGATGAATGTCAGCTAC-3' HT22R as above	To PCR Kan
	HT106F 5'-GCACTCTAGAGGGATATCCCCGAATCCATG-3' HT106R 5'-CATTGTGGCAATCGGCACATCCGTAGG-3'	To PCR *gsu2212* downstream
gsu2213::kan	HT203F 5'-GCACAATATGGTCTTCGAGCGTCAGG-3' HT203R 5'-GCATGAATTCCCGGAGACGCTCATTTTCC-3'	To PCR *gsu2213* upstream
	HT177F 5'-CGACGAATTCCCTGGGATGAATGTCAGCTAC-3' HT22R as above	To PCR Kan
	HT204F 5'-GCTATCTAGAGCCAGCGTCGACTACGAG-3' HT204R 5'-CCTGCCTGGATCGGCTTGGTAATGTAG-3'	To PCR *gsu2213* downstream
gsu2214::kan	HT112F 5'-CCAGTGGCGAGGAGCCATACACC-3' HT112R 5'-GCTAGTCGACGACCCTGATCTTTCTCATGG-3'	To PCR *gsu2214* upstream
	HT22F 5'-GCTAGTCGACTGGGATGAATGTCAGCTAC-3' HT22R as above	To PCR Kan
	HT107F 5'-GCACTCTAGACTTCTCAGCGACGTCGATTG-3' HT107R 5'-CGAACTCCTCGGCACCAATCAGG-3'	To PCR *gsu2214* downstream

gsu2215::kan	HT145F 5'-GTCACTTGACGCTATCGAGCGGCTG-3' HT145R 5'-CGTACAAGCTTCGAAGAAGAGACCGCAGTG-3'	To PCR *gsu2215* upstream
	HT42F and HT22R as above	To PCR Kan
	HT146F 5'-GGTATCTAGAGATTCCTTCCACCGCGTCC-3' HT146R 5'-GGATGGTGAGAATCTCCTCGGACACC-3'	To PCR *gsu2215* downstream
gsu2215:spec	HT139F 5'-CGAGTAAGCTTAGCACAGGATGACGCCTAAC-3' HT26R 5'- GCTCTCTAGAGCATAGTCTCCCCAGCTCTC -3'	To PCR Spec
gsu2216::kan	HT201F 5'-GCTTGCCCGGCTCATTGTTTCG-3' HT201R 5'-GCATGAATTCGATCCCCTCGATGGTTCC-3'	To PCR *gsu2216* upstream
	HT177F and HT22R as above	To PCR Kan
	HT202F 5'-GCTATCTAGAGGGCTTGATAGCAGCGACG-3' HT202R 5'-GTTCCTGGTCCTTCTTCCGGTTGTAGC-3'	To PCR *gsu2216* downstream
gsu2217::kan	HT96F 5'-CGAAGGATACGCATGGAAACCGACATCC-3' HT96R' 5'-GCATGTCGACGATCACATCCGCCATCAACG-3'	To PCR *gsu2217* upstream
	HT22F and HT22R as above	To PCR Kan
	HT97F' 5'-GCACTCTAGAAGGAGAGAACTCGGCCTG-3' HT97R 5'-GCAGATCGCCGATCTTGCCGAG-3'	To PCR *gsu2217* downstream
gsu2218::kan	HT118F 5'-GCTTCAGCTGGAAAAGGTGCTCGATTCAG-3' HT118R 5'-CGATAAGCTTCATGCGTATCCTTCGGGAGC-3'	To PCR *gsu2218* upstream
	HT42F and HT22R as above	To PCR Kan
	HT119F 5'-GCTCTCTAGACGTTGATGGCGGATGTGATC-3' HT109R 5'-GCTCGTCAGGATGCGATCGAGATTGACG-3'	To PCR gsu2218 downstream
gsu2219::kan	HT108F 5'-GGTGGAAGACCTTGTCTCCGATACAGAGG-3' HT108R 5'-GCTAGTCGACTTCCATCGGCATACCCTCC-3'	To PCR *gsu2219* upstream
	HT22F and HT22R as above	To PCR Kan
	HT109F 5'-GCTATCTAGAGACTTCTGGCTGCCTGACG-3' HT109R 5'-GCTCGTCAGGATGCGATCGAGATTGACG-3'	To PCR *gsu2219* downstream
gsu2220::kan	HT116F 5'-CCTGGCAACCAATGCCTTGCTGG-3' HT116R 5'-GCACAAGCTTGGAGACAAGGTCTTCCACC-3'	To PCR *gsu2220* upstream
	HT42F and HT22R as above	To PCR Kan
	HT117F 5'-GCTCTCTAGAGCATTGATGTAATACCCAGGAGG-3' HT117R 5'-GGAAGCAGGCGAGCTGTATTTCCTGG-3'	To PCR *gsu2220* downstream
gsu2221::kan	HT98F 5'-CGTGTCGGGCATGGTCGACATCG-3' HT98R' 5'-GCACAAGCTTGAATTCCCGGTACATTTACC-3'	To PCR *gsu2221* upstream
	HT42F and HT22R as above	To PCR Kan
	HT99F' 5'-GCACTCTAGAGATGAGGCGAACCTGTACTG-3' HT99R 5'-GGTGGTCTCCCGGGGTTTGATGATCTC-3'	To PCR *gsu2221* downstream
gsu2223::kan	HT100F 5'-CGACCTGAACGGCGAGGTCGAG-3' HT100R' 5'-GCATGTCGACCATCGATCGCATGGTGGCAG-3'	To PCR *gsu2223* upstream
	HT22F and HT22R as above	To PCR Kan
	HT101F' 5'-GCTGTCTAGAATCGTCAAGTACACCTCCTG-3' HT101R 5'-CGTACTCCGTCAGGGAGCTGAGGAC-3'	To PCR *gsu2223* downstream
gsu2224::kan	HT147F 5'-GGTCAAGTTCTACTTCGCCACCCAAGTGG-3' HT147R 5'-GCAGTGTCGACCAACTTTTTCGAACTGCTCC-3'	To PCR *gsu2224* upstream
	HT22F and HT22R as above	To PCR Kan
	HT148F 5'-GCACGTCTAGAGCGATCATTGAAGCCAAGC-3' HT148R 5'-CCTGTAATTGTCGCTGGTGCGGACAAAG-3'	To PCR *gsu2224* downstream
gsu0583::kan	HT235F 5'-GGAGGGTCATCAAGGTTTCCGAGG-3' HT235R 5'-CCAGTAGCTGACATTCATCCCCATGGCAGATGCCAGATTG-3'	To PCR *gsu0583* upstream

	HT236F 5'-CAATCTGGCATCTGCCATGGGGATGAATGTCAGCTACTGG-3' HT236R 5'-CGTCACACCTTCACTGGAGCGCGGTGGAATCGAAATCTCG-3'	To PCR Kan
	HT237F 5'-CGAGATTTCGATTCCACCGCGCTCCAGTGAAGGTGTGACG-3' HT237R 5'-CCTTCCCATAACCTCCGCTATTCGCG-3'	To PCR *gsu0583* downstream
gsu0766::kan	HT238F 5'-CATGACTGTTCCTCCTTCAGAGAGTGC-3' HT238R 5'-CCAGTAGCTGACATTCATCCCATGCTCCTCGAATTCGTCC-3'	To PCR *gsu0766* upstream
	HT239F 5'-GGACGAATTCGAGGAGCATGGGATGAATGTCAGCTACTGG-3' HT239R 5'-GGAACTGGCCGATCATGTCGCGGTGGAATCGAAATCTCG-3'	To PCR Kan
	HT240F 5'-CGAGATTTCGATTCCACCGCGACATGATCGGCCAGTTCC-3' HT240R 5'-CTCTGATTGCTTGGCAGATTGCAGG-3'	To PCR *gsu0766* downstream
gsu2372::kan	HT241F 5'-CCTCACACTGGTGCCGGTAAGTTC-3' HT241R 5'-CCAGTAGCTGACATTCATCCCATCATCGCCTCCAGCAGG-3'	To PCR *gsu2372* upstream
	HT242F 5'-CCTGCTGGAGGCGATGATGGGATGAATGTCAGCTACTGG-3' HT242R 5'-GTGATGTTCTCGGTGGAACGCGGTGGAATCGAAATCTCG-3'	To PCR Kan
	HT243F 5'-CGAGATTTCGATTCCACCGCGTTCCACCGAGAACATCAC-3' HT243R 5'-GGAACGCCAAGTTCCACGATGTCG-3'	To PCR *gsu2372* downstream
gsu1704::kan	HT132F 5'-GGAGACTGCCTCGACATGGATGTTCG-3' HT132R 5'-GCATGTCGACTTCCATGCGGATATCCCTCC-3'	To PCR *gsu1704* upstream
	HT22F and HT22R as above	To PCR Kan
	HT133F 5'-GCACTCTAGAGATGAAGAAGATAGCCAGCTG-3' HT133R 5'- CGTAACCACCAGAGCCTTTGGTCTGG-3'	To PCR *gsu1704* downstream
gsu2225::kan	HT244F 5'-CGAGATCGTGCGGGAAAAGGTCTTTCG-3' HT244R 5' CCAGTAGCTGACATTCATCCGAAACGGGTGACCGTTTCG-3'	To PCR *gsu2225* upstream
	HT245F 5'-CGAAACGGTCACCCGTTTCGGATGAATGTCAGCTACTGG-3' HT245R 5'-GGTGATGACTCTCCACCGCGGTGGAATCGAAATCTCG-3'	To PCR Kan
	HT246F 5'-CGAGATTTCGATTCCACCG**C**GGTGGAGAGTCATCACC-3' HT246R 5'-GCAAGACACCCGATTTACGGGAAAGGCTCAC-3'	To PCR *gsu2225* downstream
Primers for PCR intergenic regions between two constitutive genes in the cluster *gsu2226-gsu2209*		
	HT207F 5'-GATTCCCAGAAGGGGATC-3' HT207R 5'-CGTCCATAAGGAAGAGGATC-3'	*gsu2226-gsu2225*
	HT205F 5'-GTGGAGAGTCATCACCATC-3' HT205R 5'-GCAAGACACCCGATTTACG-3'	*gsu2225-gsu2224*

	HT188F 5'-CGATCATTGAAGCCAAGC-3' HT188R 5'-GTCGGGCATGTTGATATC-3'	*gsu2224-gsu2223*
	HT189F 5'-GATATCAACATGCCCGAC-3' HT189R 5'-GAGCAGGTTCTCCAGGTT-3'	*gsu2223-gsu2222*
	HT190F 5'-GCTTCTATGTGGTAGTGGTC-3' HT190R 5'-GGGTTGAAGATGAAGCAG-3'	*gsu2222-gsu2221*
	HT191F 5'-CTCCATTACCACCTTCAGC-3' HT191R 5'-CATCGGGAGCGAATACATC-3'	*gsu2221-gsu2220*
	HT192F 5'-CATCATTCCCGTGTTCATCC-3' HT192R 5'-TCGAGAATGATCAGGTCG-3'	*gsu2220-gsu2219*
	HT193F 5'-CGACCTGATCATTCTCGA-3' HT193R 5'-GATGACTCCTTCCACAAAGG-3'	*gsu2219-gsu2218*
	HT194F 5'-CCTTTGTGGAAGGAGTCATC-3' HT194R 5'-CGAATTTTCATGGGCCAC-3'	*gsu2218-gsu2217*
	HT195F 5'-CTTGCCCGGCTCATTGTT-3' HT195R 5'-CTCGATGGTTCCCTCATC-3'	*gsu2217-gsu2216*
	HT196F 5'-CGATTGGGTGAATGCCTATC-3' HT196R 5'-CTGGGTGAAGGCTTTGAG-3'	*gsu2216-gsu2215*
	HT197F 5'-CCTCAACCTTCTCGATACC-3' HT197R 5'-GTGGGAGAGTATTCCATCAC-3'	*gsu2215-gsu2214*
	HT198F 5'-GGAATGACGGCAGCAAAG-3' HT198R 5'-GCTATTTCCCGCTCAAGC-3'	*gsu2214-gsu2213*
	HT199F 5'-CTGCATTCCGCTGAAGAT-3' HT199R 5'-GTCCATGACCGGCATATTG-3'	*gsu2213-gsu2212*
	HT200F 5'-CAATATGCCGGTCATGGAC-3' HT200R 5'-CAAAGGGTGCAGTTGTTG-3'	*gsu2212-gsu2210*
	HT206F 5'-GGAAATCCTCCACCATGAAG-3' HT206R 5'-GGAGTAGTTGCGCACATG-3'	*gsu2210-gsu2209*
Primers for qRT-PCR		
omcZ (gsu2076)	HT216F 5'-GTCTGTAACCGCTACGGATGG-3' HT216R 5'-GTGGTGAGTATCCTGGTTGCTG-3'	
omcS (gsu2504)	HT224F 5'-GAAGAAGACCTACACCTGG-3' HT224R 5'-GTGGTGTCGGCAACATAGT-3'	
proC (gsu2541)	HT219F 5'-CCACCGATGACGATCTGTTCT-3' HT219R 5'-CATGAGCTTTTCCTCCACCAC-3'	
Primers for complementation Δ*cheR5*		
	HT167F 5'-GCAGCGAATTCGATTGGGTGAATGCCTATCG-3' HT167R 5'-GCAGCAAGCTTGACCCTGATCTTTCTCATGG-3'	To PCR *gsu2215*

APPENDIX C

MICROARRAY DATA OF Δ*CHER5* vs WT

Order	Probe Set ID	Fold change	Gene name	Annotation
1	GSU2503	11.3 up	*omcT*	cytochrome c, 6 heme-binding sites
2	GSU2504	9.4 up	*omcS*	cytochrome c, 6 heme-binding sites
3	GSU2585	8.8 up		conserved hypothetical protein
4	GSU0810	6.5 up		OmpA domain porin (beta-barrel, OmpA, OmpA)
5	GSU2501	6.3 up		cytochrome c, 6 heme-binding sites
6	GSU2502	5.8 up		spermine/spermidine synthase family protein
7	GSU3214	5.4 up		cytochrome c, 3 heme-binding sites
8	GSU2584	5.3 up		lipoprotein, putative
9	GSU0710	5.1 up		conserved hypothetical protein
10	GSU2967	5.0 up		ferritin-like domain protein
11	GSU1558	4.9 up		hypothetical protein
12	GSU0595.1	4.9 up		conserved hypothetical protein
13	GSU0713	4.7 up		hypothetical protein
14	GSU3410	4.6 up		conserved hypothetical protein
15	GSU2510	4.5 up		conserved hypothetical protein
16	GSU0919	4.5 up		conserved hypothetical protein
17	GSU3351	4.4 up		conserved hypothetical protein
18	GSU0139	4.2 up		conserved hypothetical protein
19	GSU0911	4.1 up		iron-sulfur cluster-binding protein
20	GSU0712	4.1 up		hypothetical protein
21	GSU0672	4.1 up		conserved hypothetical protein
22	GSU0711	4.1 up		endonuclease/exonuclease/phosphatase family protein
23	GSU0714	4.1 up		hypothetical protein
24	GSU2810	4.1 up		hypothetical protein
25	GSU0910	4.0 up		aldehyde:ferredoxin oxidoreductase, tungsten-containing
26	GSU2583	4.0 up	*ycaC*	isochorismatase family protein YcaC
27	GSU2586	4.0 up		hypothetical protein
28	GSU0593	4.0 up		cytochrome b, putative
29	GSU2735	4.0 up		transcriptional regulator, TetR family
30	GSU2497	3.8 up		lipoprotein, putative
31	GSU1018	3.8 up		hypothetical protein
32	GSU3289	3.8 up		ferritin-like domain protein

33	GSU0594	3.7 up		cytochrome c, 7 heme-binding sites
34	GSU2648.1	3.6 up		conserved hypothetical protein
35	GSU2735	3.6 up		transcriptional regulator, TetR family
36	GSU1559	3.6 up		hypothetical protein
37	GSU2675	3.6 up		C1 peptidase family protein
38	GSU3409	3.6 up		conserved hypothetical protein
39	GSU2811	3.6 up		cytochrome c, 2 heme-binding sites
40	GSU2507	3.5 up		sensor histidine kinase (Cache, HAMP, HisKA-HATPase_c)
41	GSU1727	3.5 up		zinc finger transcriptional regulator, TraR/DksA family
42	GSU2509	3.4 up		glycosyl transferase, group 2 family protein
43	GSU0216	3.4 up		conserved hypothetical protein
44	GSU1557	3.4 up		mechanosensitive ion channel family protein
45	GSU0803	3.4 up	*ppsA*	phosphoenolpyruvate synthase
46	GSU0539	3.4 up		conserved hypothetical protein
47	GSU2731	3.4 up	*omcC*	membrane-associated cytochrome c, 12 heme-binding sites
48	GSU3364	3.4 up		transcriptional regulator, CopG family
49	GSU1024	3.3 up	*ppcD*	cytochrome c, 3 heme-binding sites
50	GSU0941	3.2 up		sensor histidine kinase (HisKA, HATPase_c)
51	GSU1947	3.2 up		hypothetical protein
52	GSU3342	3.2 up		conserved hypothetical protein
53	GSU2495	3.2 up		cytochrome c, 26 heme-binding sites
54	GSU0193	3.1 up		L-sorbosone dehydrogenase, putative
55	GSU3341	3.1 up	*prkA*	putative serine protein kinase
56	GSU2493	3.1 up		NHL repeat domain protein
57	GSU0804	3.1 up	*wrbA*	NAD(P)H:quinone oxidoreductase flavoprotein WrbA
58	GSU1007	3.1 up		GAF domain/HD domain protein
59	GSU2498	3.1 up		lipoprotein, putative
60	GSU2788	3.1 up		conserved hypothetical protein
61	GSU2508	3.1 up		TPR domain protein
62	GSU0597	3.1 up		conserved hypothetical protein
63	GSU0920.1	3.1 up		conserved hypothetical protein
64	GSU2655.1	3.0 up		hypothetical protein
65	GSU1728	3.0 up		radical SAM domain protein, putative
66	GSU0767	3.0 up		putative porin
67	GSU0909	3.0 up		pyridine nucleotide-disulphide oxidoreductase family protein
68	GSU0802	3.0 up		oxidoreductase, short chain dehydrogenase/reductase family

69	GSU0596	3.0 up		response receiver (REC)
70	GSU2792	3.0 up		conserved hypothetical protein
71	GSU2725	2.9 up		cytochrome c, 5 heme-binding sites
72	GSU0716	2.9 up		conserved hypothetical protein
73	GSU2494	2.9 up		cytochrome c, 10 heme-binding sites
74	GSU2496	2.9 up		conserved hypothetical protein
75	GSU2649	2.9 up		amino acid ABC transporter, amino acid-binding protein
76	GSU1079	2.9 up		hypothetical protein
77	GSU2812	2.8 up		glutaredoxin family protein
78	GSU2791	2.8 up		conserved hypothetical protein
79	GSU1796.1	2.8 up		conserved hypothetical protein
80	GSU1331	2.8 up		metal ion efflux pump, RND family, membrane fusion protein
81	GSU1209	2.8 up		conserved hypothetical protein
82	GSU0195	2.8 up		protein of unknown function DUF1458
83	GSU0358	2.7 up		iron-sulfur cluster-binding protein
84	GSU0720	2.7 up		superoxide reductase-like domain (class II) protein
85	GSU1399.1	2.7 up		sensor diguanylate cyclase (PAS, GGDEF)
86	GSU3152	2.7 up		sensor protein (PAS)
87	GSU3343	2.7 up		SpoVR-like family protein
88	GSU2667	2.7 up		sensor histidine kinase (PAS, PAC, HisKA, HATPase_c)
89	GSU2813	2.7 up		cytochrome c, 2 heme-binding sites
90	GSU2732	2.7 up		cytochrome c, 8 heme-binding sites
91	GSU0200	2.6 up		aerobic-type carbon monoxide dehydrogenase, small subunit-like protein
92	GSU0201	2.6 up		aerobic-type carbon monoxide dehydrogenase, large subunit-like protein
93	GSU0715	2.6 up		conserved hypothetical protein
94	GSU0718	2.6 up		sensor histidine kinase response receiver (PAS, HisKA, HATPase_c, REC)
95	GSU0915	2.6 up		conserved hypothetical protein
96	GSU1039	2.6 up		sigma-54-dependent sensor transcriptional response regulator (REC, PAS-like, sigma54 interaction, HTH8)
97	GSU2733	2.6 up		conserved hypothetical protein
98	GSU1213	2.5 up		conserved hypothetical protein
99	GSU0717	2.5 up		conserved hypothetical protein
100	GSU1399	2.5 up	*corA-1*	magnesium and cobalt transport protein CorA
101	GSU1333	2.5 up		conserved hypothetical protein

102	GSU0067	2.5 up	*can-1*	carbonic anhydrase, beta-family, clade B
103	GSU0075	2.5 up		conserved hypothetical protein
104	GSU0907	2.5 up		ThiF family protein
105	GSU0709	2.5 up		conserved hypothetical protein
106	GSU1994	2.5 up		hypothetical protein
107	GSU2793	2.5 up		conserved hypothetical protein
108	GSU0466	2.5 up	*macA*	cytochrome c, 2 heme-binding sites
109	GSU1943	2.4 up		hypothetical protein
110	GSU2724	2.4 up		cytochrome c, 13-15 heme-binding sites
111	GSU2726	2.4 up		conserved hypothetical protein
112	GSU0078	2.4 up		PilZ domain protein
113	GSU1984	2.4 up		polysaccharide chain length determinant protein, putative
114	GSU1948	2.4 up		hypothetical protein
115	GSU3344	2.4 up		conserved hypothetical protein
116	GSU1212	2.4 up		conserved hypothetical protein
117	GSU1817	2.4 up		outer membrane lipoprotein, Slp family
118	GSU2822	2.4 up	*nasR*	response regulator (nitrate?) with putative antiterminator output domain (REC, ANTAR)
119	GSU1167	2.4 up		conserved hypothetical protein
120	GSU0674	2.4 up	*hcp*	iron-sulfur-oxygen hybrid cluster protein (prismane)
121	GSU0077	2.4 up		conserved hypothetical protein
122	GSU2743	2.4 up		cytochrome c, 1 heme-binding site
123	GSU0070	2.3 up		oxidoreductase, membrane subunit
124	GSU3357	2.3 up		sensor histidine kinase (PAS, HisKA, HATPase_c)
125	GSU1058	2.3 up	*sucC*	succinyl-CoA synthetase, beta subunit
126	GSU0194	2.3 up		conserved hypothetical protein
127	GSU1447	2.3 up		conserved hypothetical protein
128	GSU1332	2.3 up		metal ion efflux pump, RND family, inner membrane protein
129	GSU0908	2.3 up		moaD family protein
130	GSU2193	2.3 up		ferritin-like domain protein
131	GSU0700	2.3 up		response receiver sensor phosphatase (REC, PAS, PAC, PAS, PAC, PP2C)
132	GSU1404	2.2 up		radical SAM domain protein
133	GSU1330	2.2 up		metal ion efflux pump, RND family, outer membrane protein
134	GSU1995	2.2 up		conserved hypothetical protein
135	GSU0224	2.2 up		conserved hypothetical protein
136	GSU2882	2.2 up	*omcG*	cytochrome c, 14-18 heme-binding sites
137	GSU2412	2.2 up		conserved hypothetical protein

138	GSU3251	2.2 up		DUF523-containing protein
139	GSU1945	2.2 up		fibronectin type III domain protein
140	GSU1414	2.2 up		sensor histidine kinase response regulator (PAS, PAC, HisKA-HATPase_c, REC)
141	GSU2536	2.2 up		dienelactone hydrolase family protein
142	GSU2964	2.2 up	*modE*	molybdenum transport regulatory protein ModE
143	GSU3014	2.2 up		predicted signal transduction protein
144	GSU0217	2.2 up		nitroreductase 3 family protein
145	GSU3352	2.2 up		hypothetical protein
146	GSU0912	2.2 up		conserved hypothetical protein
147	GSU0719	2.2 up		conserved hypothetical protein
148	GSU1415	2.2 up		response regulator, putative (REC)
149	GSU1726	2.2 up		conserved hypothetical protein
150	GSU0071	2.1 up		conserved hypothetical protein
151	GSU3345	2.1 up		hypothetical protein
152	GSU0746	2.1 up		cytochrome p460, 1 heme-binding site
153	GSU0068	2.1 up		cytochrome c, 4 heme-binding sites
154	GSU2748	2.1 up		possible cytochrome c, 1 heme-binding site
155	GSU1448	2.1 up		metal-dependent phosphoesterase, PHP family
156	GSU1949	2.1 up		hypothetical protein
157	GSU2742	2.1 up		conserved hypothetical protein
158	GSU2749	2.1 up		NOL1/NOP2/Sun (tRNA and rRNA cytosine-C5-methylase) family protein
159	GSU0065.1	2.1 up		conserved hypothetical protein
160	GSU0726	2.0 up	*cheD-1*	chemotaxis protein CheD, putative
161	GSU0357	2.0 up		cytochrome c, 7-8 heme-binding sites
162	GSU2746	2.0 up		conserved domain protein
163	GSU2215	8.0 down	*cheR5*	chemotaxis protein methyltransferase CheR
164	GSU3388	4.1 down		membrane protein, putative
165	GSU0491	3.0 down	*rhlE-1*	ATP-dependent RNA helicase RhlE
166	GSU2886.1	3.0 down		cytochrome c, 7 heme-binding sites
167	GSU0781	2.9 down		twin-arginine translocation protein, TatA/E family
168	GSU2075	2.8 down		serine protease, subtilase family
169	GSU2517	2.8 down		rhodanese-like domain/cysteine-rich domain protein
170	GSU2887	2.7 down		cytochrome c, 27 heme-binding sites
171	GSU2074	2.6 down		PPIC-type PPIASE domain protein
172	GSU2518	2.6 down		conserved hypothetical protein
173	GSU2944	2.3 down		(R)-2-hydroxyglutaryl-CoA dehydratase D component-related protein

174	GSU0778	2.3 down	*fdnH*	formate dehydrogenase, iron-sulfur subunit
175	GSU2076	2.3 down	*omcZ*	cytochrome c, 7-8 heme-binding sites
176	GSU0780	2.2 down		formate dehydrogenase accessory protein FdhD

BIBLIOGRAPHY

Adler, J., (1975) Chemotaxis in bacteria. *Annual Review of Biochemistry* **44**: 341-356.

Alexander, R. P. & I. B. Zhulin, (2007) Evolutionary genomics reveals conserved structural determinants of signaling and adaptation in microbial chemoreceptors. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 2885-2890.

Alley, M. R. K., J. R. Maddock & L. Shapiro, (1992) Polar localization of bacterial chemoreceptor *Genes & Development* **6**: 825-836.

Alm, R. A. & J. S. Mattick, (1995) Identification of a Gene, Pilv, Required for Type-4 Fimbrial Biogenesis in Pseudomonas-Aeruginosa, Whose Product Possesses a Pre-Pilin-Like Leader Sequence. *Molecular Microbiology* **16**: 485-496.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. Zhang, W. Miller & D. J. Lipman, (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389-3402.

Anderson, R. T., H. A. Vrionis, I. Ortiz-Bernad, C. T. Resch, P. E. Long, R. Dayvault, K. Karp, S. Marutzky, D. R. Metzler, A. Peacock, D. C. White, M. Lowe & D. R. Lovley, (2003) Stimulating the in situ activity of Geobacter species to remove uranium from the groundwater of a uranium-contaminated aquifer. *Applied And Environmental Microbiology* **69**: 5884-5891.

Antommattei, F. M., J. B. Munzner & R. M. Weis, (2004) Ligand-specific activation of Escherichia coli chemoreceptor transmethylation. *Journal of Bacteriology* **186**: 7556-7563.

Antommattei, M. F. & R. M. Weis, (2006) Reversible methylation of glutamate residues in the receptor proteins of bacterial sensory systems. In: The Enzymes. F. Tamanoi & S. Clarke (eds). Elsevier, pp. 325-382.

Aravind, L. & E. V. Koonin, (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends in Biochemical Sciences* **23**: 469-472.

Aravind, L. & C. P. Ponting, (1997) The GAF domain: an evolutionary link between diverse phototransducing proteins. *Trends In Biochemical Sciences* **22**: 458-459.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, (1999) Current protocols in molecular biology. In. New York: John Wiley & Sons, Inc., pp.

Barnakov, A. N., L. A. Barnakova & G. L. Hazelbauer, (2001) Location of the Receptor-interaction Site on CheB, the Methylesterase Response Regulator of Bacterial Chemotaxis. *Journal of Biological Chemistry* **276**: 32984-32989.

Barrios, H., B. Valderrama & E. Morett, (1999) Compilation and analysis of sigma 54-dependent promoter sequences. *Nucleic Acids Research* **27**: 4305-4313.

Berg, H. C., (2000) Motile behavior of bacteria. *Physics today* **53**: 24-29.

Berg, H. C., (2003) The rotary motor of bacterial flagella. *Annual Review of Biochemistry* **72**: 19-54.

Berleman, J. E. & C. E. Bauer, (2005a) A che-like signal transduction cascade involved in controlling flagella biosynthesis in Rhodospirillum centenum. *Molecular Microbiology* **55**: 1390-1402.

Berleman, J. E. & C. E. Bauer, (2005b) Involvement of a Che-like signal transduction cascade in regulating cyst cell development in Rhodospirillum centenum. *Molecular Microbiology* **56**: 1457-1466.

Berleman, J. E., B. M. Hasselbring & C. E. Bauer, (2004) Hypercyst mutants in *Rhodospirillum centenum* identify regulatory loci involved in cyst cell differentiation. *Journal of Bacteriology* **186**: 5834-5841.

Bhaya, D., A. Takahashi & A. R. Grossman, (2001) Light regulation of type IV pilus-dependent motility by chemosensor-like elements in *Synechocystis* PCC6803. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 7540-7545.

Bhushan, B., A. Halasz, S. Thiboutot, G. Ampleman & J. Hawari, (2004) Chemotaxis-mediated biodegradation of cyclic nitramine explosives RDX9 HMX, and CL-20 by Clostridium sp EDB2. *Biochemical and Biophysical Research Communications* **316**: 816-821.

Bible, A. N., B. B. Stephens, D. R. Ortega, Z. H. Xie & G. Alexandre, (2008) Function of a chemotaxis-like signal transduction pathway in modulating motility, cell clumping, and cell length in the alphaproteobacterium *Azospirillum brasilense*. *Journal of Bacteriology* **190**: 6365-6375.

Black, W. P., Q. Xu & Z. M. Yang, (2006) Type IV pili function upstream of the Dif chemotaxis pathway in Myxococcus xanthus EPS regulation. *Molecular Microbiology* **61**: 447-456.

Black, W. P. & Z. M. Yang, (2004) *Myxococcus xanthus* chemotaxis homologs DifD and DifG negatively regulate fibril polysaccharide production. *Journal of Bacteriology* **186**: 1001-1008.

Blackhart, B. D. & D. R. Zusman, (1985) Frizzy Genes Of *Myxococcus-Xanthus* Are Involved In Control Of Frequency Of Reversal Of Gliding Motility. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **82**: 8767-8770.

Blair, D. F., (1995) How bacteria sense and swim *Annual Review Of Microbiology* **49**: 489-522.

Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. ColladoVides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau & Y. Shao, (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-&.

Bond, D. R., D. E. Holmes, L. M. Tender & D. R. Lovley, (2002) Electrode-reducing microorganisms that harvest energy from marine sediments. *Science* **295**: 483-485.

Bond, D. R. & D. R. Lovley, (2003) Electricity production by Geobacter sulfurreducens attached to electrodes. *Applied and Environmental Microbiology* **69**: 1548-1555.

Bonner, P. J., Q. Xu, W. P. Black, Z. Li, Z. M. Yang & L. J. Shimkets, (2005) The Dif chemosensory pathway is directly involved in phosphatidylethanolamine sensory transduction in Myxococcus xanthus. *Molecular Microbiology* **57**: 1499-1508.

Bren, A. & M. Eisenbach, (2000) How signals are heard during bacterial chemotaxis: Protein-protein interactions in sensory signal propagation. *Journal of Bacteriology* **182**: 6865-6873.

Briegel, A., H. J. Ding, Z. Li, J. Werner, Z. Gitai, D. P. Dias, R. B. Jensen & G. J. Jensen, (2008) Location and architecture of the Caulobacter crescentus chemoreceptor array. *Molecular Microbiology* **69**: 30-41.

Butler, S. M. & A. Camilli, (2005) Going against the grain: Chemotaxis and infection in Vibrio cholerae. *Nature Reviews Microbiology* **3**: 611-620.

Caccavo, F., D. J. Lonergan, D. R. Lovley, M. Davis, J. F. Stolz & M. J. McInerney, (1994a) Geobacter Sulfurreducens Sp-Nov, a Hydrogen-Oxidizing and Acetate-Oxidizing Dissimilatory Metal-Reducing Microorganism. *Applied and Environmental Microbiology* **60**: 3752-3759.

Caccavo, F., D. J. Lonergan, D. R. Lovley, M. Davis, J. F. Stolz & M. J. McInerney, (1994b) Geobacter sulfurreducens sp-nov, a hydrogen oxidizing and acetate oxidizing dissimilatory metal reducing organism *Applied And Environmental Microbiology* **60**: 3752-3759.

Caldon, C. E. & P. E. March, (2003) Function of the universally conserved bacterial GTPases. *Current Opinion in Microbiology* **6**: 135-139.

Campbell, E. L., K. D. Hagen, M. F. Cohen, M. L. Summers & J. C. Meeks, (1996) The devR gene product is characteristic of receivers of two-component regulatory systems and is essential for heterocyst development in the filamentous Cyanobacterium Nostoc sp strain ATCC 29133. *Journal of Bacteriology* **178**: 2037-2043.

Cantwell, B. J., R. R. Draheim, R. B. Weart, C. Nguyen, R. C. Stewart & M. D. Manson, (2003) CheZ phosphatase localizes to chemoreceptor patches via CheA-short. *Journal of Bacteriology* **185**: 2354-2361.

Chang, B. Y. & M. Dworkin, (1994) Isolated Fibrils Rescue Cohesion and Development in the Dsp Mutant of Myxococcus-Xanthus. *Journal of Bacteriology* **176**: 7190-7196.

Chao, X. J., T. J. Muff, S. Y. Park, S. Zhang, A. M. Pollard, G. W. Ordal, A. M. Bilwes & B. R. Crane, (2006) A receptor-modifying deamidase in complex with a signaling phosphatase reveals reciprocal regulation. *Cell* **124**: 561-571.

Charon, N. W. & S. F. Goldstein, (2002) Genetics of motility and chemotaxis of a fascinating group of bacteria: The spirochetes. *Annual Review Of Genetics* **36**: 47-73.

Chervitz, S. A., C. M. Lin & J. J. Falke, (1995) Transmembrane signaling by the aspartate receptor-engineered disulfides reveal static regions of the subunit interface *Biochemistry* **34**: 9722-9733.

Chevance, F. F. V. & K. T. Hughes, (2008) Coordinating assembly of a bacterial macromolecular machine. *Nature Reviews Microbiology* **6**: 455-465.

Childers, S. E., S. Ciufo & D. R. Lovley, (2002) Geobacter metallireducens accesses insoluble Fe(III) oxide by chemotaxis. *Nature* **416**: 767-769.

Chung, Y. H., M. S. Cho, Y. J. Moon, J. S. Choi, Y. C. Yoo, Y. I. Park, K. M. Lee, K. W. Kang & Y. M. Park, (2001) ctr1, a gene involved in a signal transduction pathway of the gliding motility in the cyanobacterium Synechocystis sp PCC 6803. *Febs Letters* **492**: 33-+.

Claros, M. G., and von Heijne, G. , (1994) TopPred II: An Improved Software For Membrane Protein Structure Predictions *CABIOS* **10**: 685-686.

Coppi, M. V., C. Leang, S. J. Sandler & D. R. Lovley, (2001) Development of a genetic system for Geobacter sulfurreducens. *Applied and Environmental Microbiology* **67**: 3180-3187.

Curtis, P. D., J. Atwood, R. Orlando & L. J. Shimkets, (2007) Proteins associated with the Myxococcus xanthus extracellular matrix. *Journal of Bacteriology* **189**: 7634-7642.

Darzins, A., (1994) Characterization of a Pseudomonas-Aeruginosa Gene-Cluster Involved in Pilus Biosynthesis and Twitching Motility - Sequence Similarity to the Chemotaxis Proteins of Enterics and the Gliding Bacterium Myxococcus-Xanthus. *Molecular Microbiology* **11**: 137-153.

Debabov, V. G., (2008) Electricity from microorganisms. *Microbiology* **77**: 123-131.

Djordjevic, S. & A. M. Stock, (1998) Chemotaxis receptor recognition by protein methyltransferase CheR. *Nature Structural Biology* **5**: 446-450.

Fabret, C., V. A. Feher & J. A. Hoch, (1999) Two-component signal transduction in Bacillus subtilis: How one organism sees its world. *Journal of Bacteriology* **181**: 1975-1983.

Falke, J. J., R. B. Bass, S. L. Butler, S. A. Chervitz & M. A. Danielson, (1997) The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annual Review of Cell and Developmental Biology* **13**: 457.

Falke, J. J. & G. L. Hazelbauer, (2001) Transmembrane signaling in bacterial chemoreceptors. *Trends In Biochemical Sciences* **26**: 257-265.

Felsenstein, J., PHYLIP, programs for interferring phylogenies. In.: University of Washington, pp.

Felsenstein, J., (1989) Phylogeny Interference Package (Version 3.2). *Cladistics* **5**: 164-166.

Ferrandez, A., A. C. Hawkins, D. T. Summerfield & C. S. Harwood, (2002) Cluster II che genes from *Pseudomonas aeruginosa* are required for an optimal chemotactic response. *Journal of Bacteriology* **184**: 4374-4383.

Finn, R. D., J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. L. Sonnhammer & A. Bateman, (2006) Pfam: clans, web tools and services. *Nucleic Acids Research* **34**: D247-D251.

Francis, R. T. & R. R. Becker, (1984) Specific Indication of Hemoproteins in Polyacrylamide Gels Using a Double-Staining Process. *Analytical Biochemistry* **136**: 509-514.

Galperin, M. Y., (2005) A census of membrane-bound and intracellular signal transduction proteins in bacteria: Bacterial IQ, extroverts and introverts. *BMC Microbiology* **5**.

Galperin, M. Y., (2006) Structural classification of bacterial response regulators: Diversity of output domains and domain combinations. *Journal of Bacteriology* **188**: 4169-4182.

Gardy, J. L., M. R. Laird, F. Chen, S. Rey, C. J. Walsh, M. Ester & F. S. L. Brinkman, (2005) PSORTb v.2.0: Expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics* **21**: 617-623.

Gordillo, F., F. P. Chavez & C. A. Jerez, (2007) Motility and chemotaxis of Pseudomonas sp B4 towards polychlorobiphenyls and chlorobenzoates. *Fems Microbiology Ecology* **60**: 322-328.

Gregory, K. B. & D. R. Lovley, (2005) Remediation and recovery of uranium from contaminated subsurface environments with electrodes. *Environmental Science & Technology* **39**: 8943-8947.

Grimm, A. C. & C. S. Harwood, (1997) Chemotaxis of Pseudomonas spp. to the polyaromatic hydrocarbon naphthalene. *Applied And Environmental Microbiology* **63**: 4111-4115.

Guvener, Z. T. & C. S. Harwood, (2007) Subcellular location characteristics of the Pseudomonas aeruginosa GGDEF protein, WspR, indicate that it produces cyclic-di-GMP in response to growth on surfaces. *Molecular Microbiology* **66**: 1459-1473.

Guvener, Z. T., D. F. Tifrea & C. S. Harwood, (2006) Two different Pseudomonas aeruginosa chemosensory signal transduction complexes localize to cell poles and form and remould in stationary phase. *Molecular Microbiology* **61**: 106-118.

Harrison, D. M., J. Skidmore, J. P. Armitage & J. R. Maddock, (1999) Localization and environmental regulation of MCP-like proteins in Rhodobacter sphaeroides. *Molecular Microbiology* **31**: 885-892.

Hickman, J. W., D. F. Tifrea & C. S. Harwood, (2005) A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **102**: 14422-14427.

Holmes, D. E., S. K. Chaudhuri, K. P. Nevin, T. Mehta, B. A. Methe, A. Liu, J. E. Ward, T. L. Woodard, J. Webster & D. R. Lovley, (2006) Microarray and genetic analysis of electron transfer to electrodes in Geobacter sulfurreducens. *Environmental Microbiology* **8**: 1805-1815.

Holmes, D. E., K. P. Nevin & D. R. Lovley, (2004) In situ expression of nifD in Geobacteraceae in subsurface sediments. *Applied and Environmental Microbiology* **70**: 7251-7259.

Hurley, J. H., (2003) GAF domains: Cyclic nucleotides come full circle. *Science's STKE*: PE1.

Imae, Y., K. Oosawa, T. Mizuno, M. Kihara & R. M. Macnab, (1987) Phenol - A complex chemoeffector in bacterial chemotaxis. *Journal of Bacteriology* **169**: 371-379.

Jiang, Z. Y. & C. E. Bauer, (1997) Analysis of a chemotaxis operon from Rhodospirillum centenum. *Journal of Bacteriology* **179**: 5712-5719.

Jiang, Z. Y., H. Gest & C. E. Bauer, (1997) Chemosensory and photosensory perception in purple photosynthetic bacteria utilize common signal transduction components. *Journal of Bacteriology* **179**: 5720-5727.

Kato, J., T. Nakamura, A. Kuroda & H. Ohtake, (1999) Cloning and characterization of chemotaxis genes in Pseudomonas aeruginosa. *Bioscience Biotechnology And Biochemistry* **63**: 155-161.

Kim, B. C., C. Leang, Y. H. R. Ding, R. H. Glaven, M. V. Coppi & D. R. Lovley, (2005) OmcF, a putative c-type monoheme outer membrane cytochrome required for the expression of other outer membrane cytochromes in Geobacter sulfurreducens. *Journal of Bacteriology* **187**: 4505-4513.

Kim, K. K., H. Yokota & S. H. Kim, (1999) Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature* **400**: 787-792.

Kimura, Y., S. Ishida, H. Matoba & N. Okahisa, (2004) A Myxococcus xanthus rppA-mmrA double mutant exhibits reduced uptake of amino acids and tolerance of some antimicrobials. *Fems Microbiology Letters* **238**: 145-150.

Kirby, J. R. & D. R. Zusman, (2003) Chemosensory regulation of developmental gene expression in Myxococcus xanthus. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **100**: 2008-2013.

Kirov, S. M., M. Castrisios & J. G. Shaw, (2004) Aeromonas flagella (polar and lateral) are enterocyte adhesins that contribute to biofilm formation on surfaces. *Infection and Immunity* **72**: 1939-1945.

Kort, E. N., M. F. Goy, S. H. Larsen & J. Adler, (1975) Methylation Of A Membrane Protein Involved In Bacterial Chemotaxis. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **72**: 3939-3943.

Krogh, A., B. Larsson, G. von Heijne & E. L. L. Sonnhammer, (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology* **305**: 567-580.

Krushkal, J., M. Puljic, B. Yan, J. F. Barbe, R. Mahadevan, B. Postier, R. A. O'Neil, G. Reguera, C. Leang, L. N. DiDonato, C. Nunez, B. A. Methe, R. M. Adkins, D. R. Lovley & Ieee, (2008) Genome regions involved in multiple regulatory pathways identified using GSEL, a genome-wide database of regulatory sequence elements of Geobacter sulfurreducens. In: 1st International Conference on Biomedical Engineering and Informatics. Sanya, PEOPLES R CHINA: Ieee Computer Soc, pp. 424-431.

Krushkal, J., B. Yan, L. N. DiDonato, M. Puljic, K. P. Nevin, T. L. Woodard, R. M. Adkins, B. A. Methe & D. R. Lovley, (2007) Genome-wide expression profiling in Geobacter sulfurreducens: identification of Fur and RpoS transcription regulatory sites in a rel(Gsu) mutant. *Functional & Integrative Genomics* **7**: 229-255.

Kutsukake, K., Y. Ohya & T. Iino, (1990) Transcriptional Analysis Of The Flagellar Regulon Of Salmonella-Typhimurium. *Journal of Bacteriology* **172**: 741-747.

Lahm, H. W. & H. Langen, (2000) Mass spectrometry: A tool for the identification of proteins separated by gels. *Electrophoresis* **21**: 2105-2114.

LeMoual, H. & D. E. Koshland, (1996) Molecular evolution of the C-terminal cytoplasmic domain of a superfamily of bacterial receptors involved in taxis. *Journal of Molecular Biology* **261**: 568-585.

LeMoual, H., T. Quang & D. E. Koshland, (1997) Methylation of the Escherichia coli chemotaxis receptors: Intra- and interdimer mechanisms. *Biochemistry* **36**: 13441-13448.

Letunic, I., L. Goodstadt, N. J. Dickens, T. Doerks, J. Schultz, R. Mott, F. Ciccarelli, R. R. Copley, C. P. Ponting & P. Bork, (2002) Recent improvements to the SMART domain-based sequence annotation resource. In., pp. 242-244.

Li, J. Y., G. Y. Li & R. M. Weis, (1997) The serine chemoreceptor from Escherichia coli is methylated through an inter-dimer process. *Biochemistry* **36**: 11851-11857.

Lovley, D. R., (2008a) The microbe electric: conversion of organic matter to electricity. *Current Opinion in Biotechnology* **19**: 564-571.

Lovley, D. R., M. J. Baedecker, D. J. Lonergan, I. M. Cozzarelli, E. J. P. Phillips & D. I. Siegel, (1989) Oxidation of aromatic contaminants coupled to microbial iron reduction *Nature* **339**: 297-300.

Lovley, D. R., S. J. Giovannoni, D. C. White, J. E. Champine, E. J. P. Phillips, Y. A. Gorby & S. Goodwin, (1993) Geobacter metallireducens Gen-Nov sp-nov, a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals *Archives of Microbiology* **159**: 336-344.

Lovley, D. R. & E. J. P. Phillips, (1988a) Novel Mode of Microbial Energy-Metabolism - Organic-Carbon Oxidation Coupled to Dissimilatory Reduction of Iron or Manganese. *Applied and Environmental Microbiology* **54**: 1472-1480.

Lovley, D. R. & E. J. P. Phillips, (1988b) Novel mode of microbial energy metabolism organic carbon oxidation coupled to disscimilatory reduction of iron or manganese. *Applied And Environmental Microbiology* **54**: 1472-1480.

Macnab, R. M., (1996) Flagella and motility. In *Escherichia coli and Salmonella* (Neidhardt, F.C., ed.). 123-145.

Maddock, J. R. & L. Shapiro, (1993a) Polar location of the chemoreceptor complex in *Escherichia-coli* cell *Science* **259**: 1717-1723.

Maddock, J. R. & L. Shapiro, (1993b) Polar Location of the Chemoreceptor Complex in the *Escherichia-Coli* Cell. *Science* **259**: 1717-1723.

Makarova, K. S., E. V. Koonin, R. Haselkorn & M. Y. Galperin, (2006) Cyanobacterial response regulator PatA contains a conserved N-terminal domain (PATAN) with an alpha-helical insertion. *Bioinformatics* **22**: 1297-1301.

Marchant, J., B. Wren & J. Ketley, (2002) Exploiting genome sequence: predictions for mechanisms of *Campylobacter* chemotaxis. *Trends in Microbiology* **10**: 155-159.

Marchler-Bauer, A., J. B. Anderson, P. F. Cherukuri, C. DeWweese-Scott, L. Y. Geer, M. Gwadz, S. Q. He, D. I. Hurwitz, J. D. Jackson, Z. X. Ke, C. J. Lanczycki, C. A. Liebert, C. L. Liu, F. Lu, G. H. Marchler, M. Mullokandov, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, R. A. Yamashita, J. J. Yin, D. C. Zhang & S. H. Bryant, (2005) CDD: a conserved domain database for protein classification. *Nucleic Acids Research* **33**: D192-D196.

Martin, A. C., M. Gould, E. Byles, M. A. J. Roberts & J. P. Armitage, (2006) Two chemosensory operons of Rhodobacter sphaeroides are regulated independently by sigma 28 and sigma 54. *Journal of Bacteriology* **188**: 7932-7940.

Martin, A. C., G. H. Wadhams & J. P. Armitage, (2001) The roles of the multiple CheW and CheA homologues in chemotaxis and in chemoreceptor localization in Rhodobacter sphaeroides. *Molecular Microbiology* **40**: 1261-1272.

Masduki, A., J. Nakamura, T. Ohga, R. Umezaki, J. Kato & H. Ohtake, (1995) Isolation and Characterization of Chemotaxis Mutants and Genes of Pseudomonas-Aeruginosa. *Journal of Bacteriology* **177**: 948-952.

McCarter, L. L., (2006) Regulation of flagella. *Current Opinion In Microbiology* **9**: 180-186.

Mehr, I. J., C. D. Long, C. D. Serkin & H. S. Seifert, (2000) A homologue of the recombination-dependent growth gene, rdgC, is involved in gonococcal pilin antigenic variation. *Genetics* **154**: 523-532.

Mehta, T., M. V. Coppi, S. E. Childers & D. R. Lovley, (2005) Outer membrane c-type cytochromes required for Fe(III) and Mn(IV) oxide reduction in Geobacter sulfurreducens. *Applied and Environmental Microbiology* **71**: 8634-8641.

Merritt, P. A., T. Danhorn & C. Fuqua, (2007) Motility and chemotaxis in Agrobacterium tumefaciens surface attachment and Biofilm formation. *Journal of Bacteriology* **189**: 8005-8014.

Methe, B. A., K. E. Nelson, J. A. Eisen, I. T. Paulsen, W. Nelson, J. F. Heidelberg, D. Wu, M. Wu, N. Ward, M. J. Beanan, R. J. Dodson, R. Madupu, L. M. Brinkac, S. C. Daugherty, R. T. DeBoy, A. S. Durkin, M. Gwinn, J. F. Kolonay, S. A. Sullivan, D. H. Haft, J. Selengut, T. M. Davidsen, N. Zafar, O. White, B. Tran, C. Romero, H. A. Forberger, J. Weidman, H. Khouri, T. V. Feldblyum, T. R. Utterback, S. E. Van Aken, D. R. Lovley & C. M. Fraser, (2003) Genome of *Geobacter sulfurreducens*: Metal reduction in subsurface environments. *Science* **302**: 1967-1969.

Morales, V. M., A. Backman & M. Bagdasarian, (1991) A Series of Wide-Host-Range Low-Copy-Number Vectors That Allow Direct Screening for Recombinants. *Gene* **97**: 39-47.

Munch, R., K. Hiller, A. Grote, M. Scheer, J. Klein, M. Schobert & D. Jahn, (2005) Virtual footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* **21**: 4187-4189.

National Research Council, (1999) *Groundwater and Soil Cleanup*. National Academy Press.

Nevin, K. P., B.-C. Kim, R. H. Glaven, J. P. Johnson, T. L. Woodard, B. A. Methe, R. J. DiDonato, Jr., S. F. Covalla, A. E. Franks, A. Liu & D. R. Lovley, (2009) Anode Biofilm Transcriptomics Reveals Outer Surface Components Essential for High Density Current Production in *Geobacter sulfurreducens* Fuel Cells. *PLoS ONE* **4**: e5628.

Nevin, K. P., H. Richter, S. F. Covalla, J. P. Johnson, T. L. Woodard, A. L. Orloff, H. Jia, M. Zhang & D. R. Lovley, (2008) Power output and columbic efficiencies from biofilms of Geobacter sulfurreducens comparable to mixed community microbial fuel cells. *Environmental Microbiology* **10**: 2505-2514.

O'Toole, G., H. B. Kaplan & R. Kolter, (2000) Biofilm formation as microbial development. *Annual Review Of Microbiology* **54**: 49-79.

Page, R. D. M., (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**: 357-358.

Park, S. Y., P. P. Borbat, G. Gonzalez-Bonet, J. Bhatnagar, A. M. Pollard, J. H. Freed, A. M. Bilwes & B. R. Crane, (2006) Reconstruction of the chemotaxis receptor-kinase assembly. *Nature Structural & Molecular Biology* **13**: 400-407.

Parkinson, J. S., P. Ames & C. A. Studdert, (2005) Collaborative signaling by bacterial chemoreceptors. *Current Opinion In Microbiology* **8**: 116-121.

Perez, E. & A. M. Stock, (2007) Characterization of the Thermotoga maritima chemotaxis methylation system that lacks pentapeptide-dependent methyltransferase CheR : MCP tethering. *Molecular Microbiology* **63**: 363-378.

Ponting, C. P. & L. Aravind, (1997) PAS: a multifunctional domain family comes to light. *Current Biology* **7**: R674-R677.

Porter, S. L., A. V. Warren, A. C. Martin & J. P. Armitage, (2002) The third chemotaxis locus of Rhodobacter sphaeroides is essential for chemotaxis. *Molecular Microbiology* **46**: 1081-1094.

Postier, B., R. DiDonato Jr, K. P. Nevin, A. Liu, B. Frank, D. Lovley & B. A. Methe, (2008) Benefits of in-situ synthesized microarrays for analysis of gene expression in understudied microorganisms. *Journal of Microbiological Methods* **74**: 26-32.

Reese, M. G., N. L. Harris & F. H. Eeckman, (1996) Large Scale Sequencing Specific Neural Networks for Promoter and Splice Site Recognition. In: Biocomputing: Proceedings of the 1996 Pacific Symposium. World Scientific Publishing Co, Singapore, pp.

Reguera, G., K. D. McCarthy, T. Mehta, J. S. Nicoll, M. T. Tuominen & D. R. Lovley, (2005) Extracellular electron transfer via microbial nanowires. *Nature* **435**: 1098-1101.

Robinson, V. L., J. Hwang, E. Fox, M. Inouye & A. M. Stock, (2002) Domain arrangement of Der, a switch protein containing two GTPase domains. *Structure* **10**: 1649-1658.

Rosario, M. M. L. & G. W. Ordal, (1996) CheC and CheD interact to regulate methylation of Bacillus subtilis methyl-accepting chemotaxis proteins. *Molecular Microbiology* **21**: 511-518.

Samanta, S. K., B. Bhushan, A. Chauhan & R. K. Jain, (2000) Chemotaxis of a Ralstonia sp SJ98 toward different nitroaromatic compounds and their degradation. *Biochemical and Biophysical Research Communications* **269**: 117-123.

Sandler, S. J. & A. J. Clark, (1994) Recor Suppression of Recf Mutant Phenotypes in Escherichia-Coli K-12. *Journal of Bacteriology* **176**: 3661-3672.

Sauer, K., M. C. Cullen, A. H. Rickard, L. A. H. Zeef, D. G. Davies & P. Gilbert, (2004) Characterization of nutrient-induced dispersion in Pseudomonas aeruginosa PAO1 biofilm. *Journal of Bacteriology* **186**: 7312-7326.

Segall, J. E., S. M. Block & H. C. Berg, (1986) Temporal comparisons in bacterial chemtoaxis *Proceedings Of The National Academy Of Sciences Of The United States Of America* **83**: 8987-8991.

Shelobolina, E. S., H. A. Vrionis, R. H. Findlay & D. R. Lovley, (2008) Geobacter uraniireducens sp nov., isolated from subsurface sediment undergoing uranium bioremediation. *International Journal of Systematic and Evolutionary Microbiology* **58**: 1075-1078.

Shrout, A. L., D. J. Montefusco & R. M. Weis, (2003) Template-directed assembly of receptor signaling complexes. *Biochemistry* **42**: 13379-13385.

Singh, R., and Olson Mira S., (2008) Application of bacterial swimming and chemotaxis for enhanced bioremediation. In: Emerging Environmental Technologies. V. Shah (ed). Springer Science, pp. 149-172.

Smith, P. K., R. I. Krohn, G. T. Hermanson, A. K. Mallia, F. H. Gartner, M. D. Provenzano, E. K. Fujimoto, N. M. Goeke, B. J. Olson & D. C. Klenk, (1985) Measurement of Protein Using Bicinchoninic Acid. *Analytical Biochemistry* **150**: 76-85.

Sourjik, V. & H. C. Berg, (2000) Localization of components of the chemotaxis machinery of Escherichia coli using fluorescent protein fusions. *Molecular Microbiology* **37**: 740-751.

Sourjik, V. & H. C. Berg, (2004) Functional interactions between receptors in bacterial chemotaxis. *Nature* **428**: 437-441.

Sourjik, V. & R. Schmitt, (1998) Phosphotransfer between CheA, CheY1, and CheY2 in the Chemotaxis Signal Transduction Chain of Rhizobium meliloti. *Biochemistry* **37**: 2327-2335.

Soutourina, O. A. & P. N. Bertin, (2003) Regulation cascade of flagellar expression in Gram-negative bacteria. *Fems Microbiology Reviews* **27**: 505-523.

Stoffel, K. H. W., (1993) TMbase - A database of membrane spanning proteins segments. *Biological Chemistry Hoppe-Seyler* **374**: 116.

Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. L. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrock-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. S. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. W. Hancock, S. Lory & M. V. Olson, (2000) Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. *Nature* **406**: 959-964.

Strycharz, S. M., T. L. Woodard, J. P. Johnson, K. P. Nevin, R. A. Sanford, F. E. Loffler & D. R. Lovley, (2008) Graphite electrode as a sole electron donor for reductive dechlorination of tetrachlorethene by Geobacter lovleyi. *Applied And Environmental Microbiology* **74**: 5943-5947.

Studholme, D. J., M. Buck & B. T. Nixon, (2000) Identification of potential sigma(N)-dependent promoters in bacterial genomes. *Microbiology-UK* **146**: 3021-3023.

Szurmant, L. & G. W. Ordal, (2004) Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiology And Molecular Biology Reviews* **68**: 301-+.

Thomas, P. E., D. Ryan & W. Levin, (1976) Improved Staining Procedure for Detection of Peroxidase-Activity of Cytochrome-P-450 on Sodium Dodecyl-Sulfate Polyacrylamide Gels. *Analytical Biochemistry* **75**: 168-176.

Thompson, J. D., D. G. Higgins & T. J. Gibson, (1994) Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research* **22**: 4673-4680.

Tran, H. T., J. Krushkal, F. M. Antommattei, D. R. Lovley & R. M. Weis, (2008) Comparative genomics of Geobacter chemotaxis genes reveals diverse signaling function. *Bmc Genomics* **9**: 15.

Vlamakis, H. C., J. R. Kirby & D. R. Zusman, (2004) The Che4 pathway of Myxococcus xanthus regulates type IV pilus-mediated motility. *Molecular Microbiology* **52**: 1799-1811.

Wadhams, G. H. & J. P. Armitage, (2004) Making sense of it all: Bacterial chemotaxis. *Nature Reviews Molecular Cell Biology* **5**: 1024-1037.

Wadhams, G. H., A. V. Warren, A. C. Martin & J. P. Armitage, (2003) Targeting of two signal transduction pathways to different regions of the bacterial cell. *Molecular Microbiology* **50**: 763-770.

Whitchurch, C. B., A. J. Leech, M. D. Young, D. Kennedy, J. L. Sargent, J. J. Bertrand, A. B. T. Semmler, A. S. Mellick, P. R. Martin, R. A. Alm, M. Hobbs, S. A. Beatson, B. Huang, L. Nguyen, J. C. Commolli, J. N. Engel, A. Darzins & J. S. Mattick, (2004) Characterization of a complex chemosensory signal transduction system which controls twitching motility in Pseudomonas aeruginosa. *Molecular Microbiology* **52**: 873-893.

Wu, J. G., J. Y. Li, G. Y. Li, D. G. Long & R. M. Weis, (1996) The receptor binding site for the methyltransferase of bacterial chemotaxis is distinct from the sites of methylation. *Biochemistry* **35**: 4984-4993.

Wuichet, K., R. P. Alexander & I. B. Zhulin, (2007) Comparative genomic and protein sequence analyses of a complex system controlling bacterial chemotaxis. In: Two-Component Signaling Systems, Pt A. pp. 3-31.

Yang, Z. M., X. Y. Ma, L. M. Tong, H. B. Kaplan, L. J. Shimkets & W. Y. Shi, (2000) Myxococcus xanthus dif genes are required for biogenesis of cell surface fibrils essential for social gliding motility. *Journal of Bacteriology* **182**: 5793-5798.

Yi, H., K. P. Nevin, B.-C. Kim, A. E. Franks, A. Klimes, L. M. Tender & D. R. Lovley, Selection of a variant of Geobacter sulfurreducens with enhanced capacity for current production in microbial fuel cells. *Biosensors and Bioelectronics* **In Press**.

Zimmerman, M. D., M. Proudfoot, A. Yakunin & W. Minor, (2008) Structural insight into the mechanism of substrate specificity and catalytic activity of an HD-domain phosphohydrolase: The 5'-deoxyribonucleotidase YfbR from Escherichia coli. *Journal of Molecular Biology* **378**: 215-226.

Zusman, D. R., A. E. Scott, Z. Yang & J. R. Kirby, (2007) Chemosensory pathways, motility and development in Myxococcus xanthus. *Nature Reviews Microbiology* **5**: 862-872.